8-2022

Three Chapters on Investments and Financial Institutions

Cao Fang
University of Arkansas, Fayetteville

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Finance

By

Cao Fang
Tianjin University
Bachelor of Science in Chemistry, 2006
Nankai University
Master of Science in Chemistry, 2009
Southern New Hampshire University
Master of Science in Finance, 2014

August 2022
University of Arkansas

This dissertation (or thesis) is approved for recommendation to the Graduate council.

Wayne Y. Lee, Ph.D. Dissertation Director

Craig G. Rennie, Ph.D.
Committee Member

Timothy J. Yeager, Ph.D.
Committee Member

Abstract

Only the stock selection ("alpha") decisions of fund managers who trade on firm-specific information should have predictive return content. Faced with the same information, skilled fund managers make similar stock selection decisions. In Chapter one, we introduce a new measure - stock investment quality - which uses fund quality to weight asymmetries in private information reflected in deviations of fund from peer group ownership on stocks in a style segment. We show stocks ranked high on investment quality generate significantly higher excess returns that persist through the ensuing year. The positive investment quality–future return relationship is robust to alternative fund quality proxies.

In Chapter two, we use fund flow shocks from exogenous changes in ETF share demand to quantify the cost of trading stocks purchased or sold by APs in conjunction with the creation or redemption of ETF shares. We document a negative relation between return and the impact of primary market activities of APs on the liquidity of ETF-owned stocks. The stock-specific liquidity effect cannot be attributed to systematic asset pricing factors. Further, we find the improvements in liquidity from the primary market activities of APs enhance price discovery and strengthen the stock return-volatility relation.

In Chapter three, we develop a top-down macro stress test that assesses a community bank's ability to withstand a severe and prolonged period of high credit losses. The model groups banks by geography and subjects them to the 90th percentile chargeoffrates that banks experienced between 2008 and 2012. Because of local data limitations, our historical loss approach better reflects patterns of community bank stress than a linear econometric approach that estimates the relationship between macroeconomic conditions and bank performance. We put all U.S. community banks at year-end 2017 through the test and highlight two results. First, banks are much better prepared to withstand an adverse shock than they were on the verge of the financial crisis because banks have

shifted away from the riskiest loan types. Second, the Tax Cuts and Jobs Act of 2017 has increased bank insolvency risk from an adverse shock in 2018 because the higher bank capital is more than offset by the weaker automatic stabilizer effect from operating losses.

Acknowledgements

I am deeply indebted to my advisor Dr. Lee for his invaluable mentorship and patience. Also, I could not have finished this journey without the generous help from my committee members Dr. Rennie and Dr. Yeager. Additionally, this journey would not have been possible without the help from Finance Department Chair, Dr. Liu.

Lastly, I would like to express my deepest gratitude to my parents. Their belief in me helps me overcome the hurdles which I could not have overcome in my life.

Table of Contents

List of Published Papers

Cao Fang, Timothy J. Yeager. 2020. A historical loss approach to community bank stress testing. Journal of Banking and Finance, 118, 105831.

Stocks through a Looking Glass:

Can Style Segment-Adjusted Mutual Fund Stock Holdings Predict Stock Returns?

Cao Fang and Wayne Y. Lee[a]

Abstract

Only the stock selection ("alpha") decisions of fund managers who trade on firm-specific information should have predictive return content. Faced with the same information, skilled fund managers make similar stock selection decisions. We introduce a new measure - stock investment quality - which uses fund quality to weight asymmetries in private information reflected in deviations of fund from peer group ownership on stocks in a style segment. We show stocks ranked high on investment quality generate significantly higher excess returns that persist through the ensuing year. The positive investment quality–future return relationship is robust to alternative fund quality proxies.

Keywords: Investment quality; Forecast stock returns; Fund quality; Active management; Turnover.

JEL classification: G10; G11; G12; G14.

[a]WLee@walton.uark.edu, 479-575-3944, Sam M. Walton College of Business, University of Arkansas, WCOB 343, Fayetteville, AR

I. Introduction

In the mutual fund literature, holdings are used to show that fund managers have skill. A large number of studies find that actively managed funds whose holdings deviate significantly from passive index benchmarks generate returns that beat their benchmarks (Cremers and Petajisto, 2009; and Petajisto, 2013). A few studies find that significant differences in the holdings of the same stock across actively managed funds forecast future stock returns (Wermers, Yao, and Zhao, 2012).[1]

Stock returns, however, have systematic and firm-specific components. In "alpha" (stock selection) strategies, fund managers trade on idiosyncratic factors, and in "beta" (asset allocation) strategies, on systematic factors (Kacperczyk, Nieuwerburgh, and Veldkamp, 2014). When fund holdings intertwine "alpha" and "beta" strategies, can the variation in holdings from a passive index benchmark meaningfully uncover the stock selection skills of fund managers?[2] Can fund-stock variation not associated with trading on firm-specific information have predictive stock return content? Last but not least, will comparably informed fund managers make similar stock selection decisions when fund managers are not equally skilled and the differences in skills across fund managers are considerable (Pastor, Stambaugh and Taylor, 2015)?

> "(i)t is far better to weight the … *informed* … opinions of more capable decision makers more heavily than those of less capable decision makers. … (B)est decisions are made by an idea meritocracy with believability-weighted decision making." Ray Dalio, 2019[3]

In this study, we introduce a stock investment quality (IQ) measure which signals future stock

---

[1]See also Chen, Hong and Stein (2002), Jiang and Sun (2014), Jiang, Verbeek and Wang (2014).

[2]Accounting for differences in benchmark returns, fund return outperformance from active share is a result of under-performing benchmarks (Frazzini, Friedman, and Pomorski, 2016). Inferring stock selection skill from active share (as) is predicated on the assumption that benchmark portfolios are zero-alpha. From Cremers and Petajisto (2009, p. 3335), $as = 0.5 \cdot \sum_i |w_{ip} - w_{ib}|$, where w_{ip} and w_{ib} denote the period t stockholdings of the fund and benchmark. Noting that $0 = \sum_i (w_{ip} - w_{ib})$, it can be shown $as = \sum_i max(w_{ip} - w_{ib}, 0) = \sum_i max(w_{ip}, w_{ib}) - \sum_i w_{ib}$ represents a call option. Excess portfolio return $r_{p,t+1} = \sum_i |w_{ip} - w_{ib}| \cdot r_{i,t+1}$, where $r_{i,t+1}$ denote period $t+1$ stock returns. Further, $r_{i,t+1} = \alpha_{i,t+1} + \sum_k \beta_{ik} r_{k,t+1} + \varepsilon_{i,t+1}$ where $\mathbf{r_k}$ denotes the vector of asset pricing factors. Predicted portfolio excess return $E(r_{p,t+1} | \mathbf{r_k}) = 2 \cdot \sum_i \{max(w_{ip}, w_{ib}) \cdot \alpha_{i,t+1}\}$ when $\sum_i w_{ib} \cdot \alpha_{i,t+1} = 0$, i.e., the benchmark portfolio is zero-alpha. $E(r_{p,t+1} | \mathbf{r_k})$ will correlate positively with active share, $|w_{ip} - w_{ib}|$ when funds overweight high alpha stocks and underweight low alpha stocks.

[3]Co-Chairman, Bridgewater Associates. See https://www.linkedin.com/pulse/work-principle-5-believability-weight-your-decision-making-ray-dalio.

returns. The measure extracts the private information in stock selection decisions made by active fund managers and gives more credence to the private information of active fund managers who have better historical records of outperformance. Just as in basketball, given the same opportunity to take a three-point shot, more skilled players are more likely to make the shot successfully on a repeated basis.

Our new measure of stock investment quality makes three important contributions. First, only skilled fund managers faced with the same information will act similarly (Cohen, Coval, and Pastor, 2005). We expect variations in fund-stock ownership driven by trading on firm-specific information to be positively correlated with skill, and variations in fund-stock ownership from sentiment-based trading of unskilled fund managers to be uncorrelated with skill.[4] Moreover, we expect the co-movements in fund-stock ownership among skilled fund managers to reflect the commonality in private information, and co-movements in fund-stock ownership of unskilled fund managers to reflect herding. Only the co-movements in fund-stock ownership from trading on private information by skilled fund managers will endure. Analogous to Wermers, Yao, and Zhao (2012), our stock IQ measure assigns more credence to the fund-stock ownership of high-quality funds.

Second, to disentangle timing from stock selection in fund holdings, we use stock characteristics (Daniel, Grinblatt, Titman, and Wermers, 1997) rather than passive index benchmarks to capture the beta exposures of stocks to systematic factors. Returns on stocks sorted into style segments will share similar beta exposures to a common set of systematic factors. When a typical fund holds 57-72 stocks (Shawky and Smith, 2005), passive index benchmarks will mirror beta exposures to systematic factors that have little in common with funds to which the same benchmark is assigned and may not be zero-alpha.[5]

[4] Let $h_{i,j}$ and $h_{i,j*}$ denote the holdings of stock i by funds j and j^*; α_j and α_{j*}, the fund performance of funds j and j^* which proxy for managerial skill. $Cov(h_{i,j}, \alpha_j)$ and $Cov(h_{i,j*}, \alpha_{j*})$ will embed $Cov(h_{i,j}, h_{i,j*})$ when funds j and j^* have skilled managers, and as a result, α_j and α_{j*} are correlated. Moreover, $Cov(h_{i,j}, h_{i,j*})$ is accentuated when $h_{i,j}$ and $h_{i,j*}$ are also functions of skill; that is, skilled managers of funds j and j^* act similarly on the same information.

[5] For a vast majority of funds, portfolio holdings on passive index benchmarks may not capture the perceived investment opportunities of fund managers and their true exposures to size and value/growth dimensions (Cremers, Fulkerson, and

Moreover, style segments allow us to decompose the variation in fund holdings into two principal components. The dollar allotments across stocks within a style segment will reflect trading on firm-specific information (stock selection), and dollar allotments across style segments, will reflect trading on systematic factors (asset allocation). Fund managers who invest in a style segment constitute a natural peer group whose members adopt similar unobserved stock selection strategies (Hunter, Kandel, Kandel, and Wermers, 2014). We use the dollar allotments of fund and peer group across stocks in a style segment to identify stock selection and the asymmetry in firm-specific information across fund managers.

A fund's stock ownership is the fraction of total dollars invested in a style segment allocated to a stock, and peer group stock ownership, the fraction of aggregate dollars invested in a style segment allocated to a stock. By construction, deviations in fund from peer group stock ownership summed across funds is zero, and the variance in stock ownership of fund from peer group characterizes the dispersion in beliefs. "Differences in beliefs must derive ultimately from differences in information" (Black, 1986).

Third, we use gross value-added (GVA) rather than gross or net alpha to proxy fund quality; that is, the skill of fund managers. Fund GVA accounts for diminishing returns to skill, the competition for assets, and the constrained supply of skilled fund managers (Berk and Green, 2004). Successful performing funds will employ more skilled managers and skillful fund managers will choose to join competitive fund families where performance incentives are high (Evans, Prado, and Zambrana, 2020). The distribution of gross value-added will predominantly mirror the distribution of managerial skill (Berk and Binsbergen, 2015) as the size of the mutual fund industry changes with the entry of skilled and exit of unskilled funds (Pastor, Stambaugh, and Taylor, 2015). When investors can identify

Riley, 2019). Whereas a small fund can easily invest all its money in its best ideas, a lack of liquidity can force a large fund to invest in its not-so-good ideas or take larger ownership positions in stocks than is optimal for risk diversification (Chen, Hong, Huang, and Kubrick, 2004). Diseconomies of fund size can conceal managerial skill (Zhu, 2018). The empirical relation between fund size and performance is an outcome of investment style and liquidity (Yan, 2008).

talent, more skilled fund managers earn higher economic rents, manage more assets, and asset prices are more information efficient (Gârleanu and Pedersen, 2018).

We compute stock IQ as the summed cross-product of fund quality and deviations of fund from peer group stock ownership scaled by the variance in fund from peer group stock ownership across funds. Stock IQ is higher when managers who are better informed are also skilled; that is, when deviations of fund from peer group stock ownership are positively correlated with fund quality (Cohen, Coval, and Pastor, 2005). The stock return signals contained in stock IQ is a "looking glass" on the collective wisdom and skill of active fund managers in forecasting future stock returns.

Our main findings are as follows. We show our stock IQ measure strongly persists up to lead four quarters. Using horserace regressions, we find the forecast return power of stock IQ is not subsumed by alternative empirically documented stock-return prediction measures which include herding by unskilled fund managers (Lakonishok, Shleifer, and Vishny, 1992; Jones, Lee, and Weis, 1999; Brown, Wei, and Wermers, 2014), adjusted 'dumb' money inflows (Frazzini and Lamont, 2007) from investor-sentiment driven trading, as well as delta fund ownership and delta breadth that may suggest short-sale constraints (Chen, Hong, and Stein, 2002). Controlling for competing stock return predictors, stock IQ significantly predicts positive stock returns over lead four quarters.

We find a strong positive relationship between stock IQ and future stock returns. A value-weighted portfolio of stocks in the highest IQ quintile outperforms a value-weighted portfolio of stocks in the lowest IQ quintile by an average quarterly excess market return of 1.533%. The buy-and-hold return outperformance of high over low IQ stocks persists up to a year.

Stock IQ signals the information advantage of skillful fund managers. Small cap stocks benefit the most from IQ, over and beyond large cap stocks. The average quarterly excess market return in lead one quarter on value-weighted high-low IQ quintile portfolios is 1.825% on small cap stocks. In comparison, it is 1.596% on midcap stocks, and 1.646% on portfolios of large cap stocks. Forecast returns are robust to alternative adjustments for risk and strongly persist through the year. Results are robust to using DGTW returns and 4-factor alphas.

Regression results confirm that forecast stock returns increase with stock IQ, and the information advantage of skilled fund managers decays slowly. Accounting for delta breadth, delta active mutual fund ownership, and dispersion in holdings as well as other controls, two-way style and quarter fixed effect regressions substantiate a significant positive investment quality–forecast return relationship. A value-weighted portfolio of stocks in the 95^{th} IQ percentile outperforms a similar portfolio of IQ stocks in the 5^{th} percentile by a quarterly excess market return of 2.019%. Forecast returns from IQ decline each quarter but strongly persist through four quarters. In the fourth quarter, forecasted average quarterly excess market return is 75% of first quarter excess market return. Results using DGTW return and 4-factor alpha are analogous. As expected, our results are both statistically and economically unchanged when we proxy fund quality by management fees or by industry concentration (Kacperczyk, Sialm, and Zheng, 2005).

Large cap stocks benefit less from IQ than small cap stocks, which should not be surprising since large cap stocks attract more attention, are more closely scrutinized, and more transparent. On long-short value-weight portfolios of stocks in the 95^{th} and 5^{th} percentiles of IQ, average forecast quarterly excess market return in lead one quarter is 2.330% on portfolios of small cap stocks. In contrast, it is 1.939% on portfolios of midcap stocks, and 1.654% on portfolios of large cap stocks. The same pattern is true for DGTW return and 4-factor alpha.

To assess the importance of conviction quality, we examine the forecast power of private information embedded in active stock ownership turnover. We define active stock ownership as the deviation of fund from peer group stock ownership and turnover as the sum product of four-quarter change in active stock ownership and fund quality. The more patient are fund managers, the lower is their active stock ownership turnover and the higher their conviction quality. Future returns are highest on active stock ownership of skillful fund managers who patiently exploit long term market mispricing. Patience is not, however, a substitute for skill. Accounting for the patience and conviction of fund managers, active stock ownership turnover does not diminish forecast quarterly returns from IQ. On stocks where active stock ownership turnover is high, however, forecast quarterly returns fall significantly.

Lastly, we examine the private information of skilled fund managers impounded in stock IQ that is made public in earnings announcements. Stock IQ strongly predicts cumulative abnormal return in the three-day windows around earnings announcements and in the post-earnings periods between quarterly earnings announcements. The post-earnings announcement drift in cumulative abnormal returns is consistent with a gradual diffusion of private fundamental information in stock prices.

In a closely related study to ours, Wermers, Yao, and Zhao (2012) use the identity that forecasted fund return is a sum product of value-based portfolio holdings and expected stock returns, to derive a "generalized inverse alpha" (GIA) measure of stock quality that efficiently extracts the private information of skilled fund managers about future stock returns from their value-weighted portfolio holdings.[6] Both their GIA and our stock IQ measure strongly predict future stock returns over one-year horizons. Sorting stocks into deciles by GIA, their Table 2 shows high-low decile portfolios generate a characteristic-adjusted return and 4-factor alpha of 1.14% and 1.15% respectively in the lead quarter: 2.60% and 3.44% respectively in the lead four quarters. Comparably, sorting stocks into quintiles by our stock IQ measure, we show in our Table 6, a high-low quintile generates a characteristic-adjusted return and 4-factor alpha of 1.723% and 1.537% respectively in the lead

[6]See equation 3. From an inverse linear projection of an $(N \times J)$ matrix of fund-portfolio value-based holdings onto a $(J \times 1)$ vector of fund characteristics-based return alphas, stock quality is an $(N \times 1)$ vector of implied consensus stock alphas computed as the matrix product of an inverse $(N \times N)$ covariance matrix of fund-portfolio value-based holdings and an $(N \times 1)$ vector of the cross-product of fund-portfolio value-based holdings and fund characteristics-based return alphas.

[6]The covariance matrix of fund-portfolio value-based holdings captures the cross-sectional variance in the dispersion of private information between skilled and unskilled fund managers.

quarter; 1.516% and 1.078% respectively in the average lead four quarters.

There are, however, important differences between GIA and our stock IQ measure. First and foremost, the inverse projection of fund portfolio holdings on fund alphas does not distinguish the within from between group variation in the portfolio holdings of skilled and unskilled fund managers. Our stock IQ recognizes the variance in the dispersion of information across skilled and unskilled fund managers, but only skilled fund managers will make informed portfolio investment decisions that are positively correlated. The sum product of active holdings and fund GVA proxy for differences in information and managerial skill.

Second, the GIA approach is novel but impractical without strong restrictions when the number of stocks held by mutual funds far exceeds the number of funds, $N \gg M$[7]. The issue of invertibility in the covariance matrix of GIA restricts the number of permissible non-zero eigenvalues K to be a sufficient order of magnitude of N such that $K < M$.[8] GIA can only be estimated for a small subset of stocks, *K*, especially in early sample years when number of funds M is small. Our stock IQ is not subject to such an estimation constraint.

Third, in other related studies, deviations of a fund's portfolio allocation to a stock from their value-weights in assigned "index" benchmarks are used to describe active share (Cremers and Petajisto, 2009), and the dispersion in active share (Cremers et al., 2009), to infer the asymmetry in information across fund managers (Jiang and Sun, 2014; and Jiang, Verbeek, and Wang, 2020). Active share will be high when fund managers who receive positive information signals can act freely to increase holdings, and low, when fund managers who receive negative information signals are subject to short-sale constraints. But as we note earlier, active share (Cremers and Petajisto, 2009) or tracking error (Cremers and Pareek, 2016) using index-based benchmarks confounds stock selection with

[7]N denotes the number of stocks in sample, and M denotes the number of funds.
[8]Wermers et al. (2012, p. 3496) set $K = M/2$ and treat the remaining $(N - K)$ eigenvalues as zero.

timing.

Jiang and Sun (2014) show a value-weighted high-low quintile portfolio spreads, sorted on quarterly changes in average standard deviation of active shares across funds, generate an average monthly DGTW return of 0.55% and 4-factor alpha of 0.49% in lead three months.[9] In multivariate regressions, predicted stock returns on quarterly changes in average dispersion persist through four quarters but rapidly deteriorates to about 14% of first quarter return in the fourth quarter.[10] In Jiang et al. (2020), the equal-weighted average of deviations of fund from benchmark holdings describes the consensus in beliefs of fund managers. A high-low quintile portfolio of stocks sorted by consensus belief generates an average monthly DGTW return of 0.38% and 4-factor alpha of 0.31% in lead three months.[11] Differential returns on high-low quintile consensus portfolios do not reverse as price pressure from abnormal demand (Gompers and Metrick, 2001) or herding behavior (Sias, 2004) suggests, but quickly converge to zero after the first quarter.[12]

Our high-low quintile portfolio of stocks sorted on stock IQ generates an average monthly DGTW return of 0.574% and 4-factor alpha of 0.512% in lead three months similar to Jiang and Sun (2014) but returns strongly persist through the ensuing year. DGTW return and 4-factor alpha in the fourth quarter are 88.0% and 70.1% of first quarter stock returns.

II. Active Performance Measures

A. Active Stock Ownership by Style Segment

At the end of July each year, we sequentially sort all common stocks into 125 (5×5×5) portfolios by size, industry-adjusted book-to-market ratio, and momentum. Size quintiles use breakpoints based on NYSE stocks. Industry-adjusted book-to-market and momentum quintiles use breakpoints based

[9]See Jiang et al. (2014) Table 3 Panel D. High-low spread on average monthly DGTW return of 0.545=0.5*(0.27+0.30) -0.5*(-0.24-0.28) and 4-factor alpha of 0.485=0.5*(0.21+0.34)-0.5*(-0.21-0.23).

[10]See Jiang and Sun (2014) Table 6 Panel B.

[11]See Jiang et al. (2020) Table 2. High-low spread on average monthly DGTW return of 0.38%=0.5*(0.35+0.32)-0.5*(0.02-0.11) and 4-factor alpha of 0.305%=0.5*(0.33+0.24)-0.5*(0.00-0.04).

[12]See Jiang et al. (2020) Figure 1.

on all stocks in each size quintile. From CRSP, size is the product of adjusted price and number of adjusted shares outstanding at June end for each firm. From S&P Compustat, book-to-market is the ratio of the book value of equity and market capitalization at the end of the fiscal year closest but prior to June each year (Daniel and Titman, 2006). To industry-adjust book-to-market, we follow Wermers (2003). The difference in the natural logs of a firm's book-to-market and average book-to-market of the industry to which the firm belongs is normed by the standard deviation of the natural log differences across firms in the industry.[13] Momentum is computed as the prior 12-month return by May end to avoid bid-ask bounce and monthly return reversals (Jegadeesh, 1990).

Using cutoffs from annual sorts, we assign the stockholdings of each fund in the subsequent four quarters into one of $k = 125$ style segments. In each quarter, we denote stock i in style segment k by $i(k)$, fund by j, and the set of stocks in style segment k owned by fund j by $i(k) \epsilon I(k,j)$. Funds who own the same stock in a style segment constitute a natural peer group. Peer group holdings capture the commonality in information and similarity in unobserved investment strategies across fund managers.

Let $h_{i,j}$ denote the percentage of total assets under management (AUM) allocated to stock i by fund j.

$$h_{i,j} = h_{k,j} h_{i(k),j} \tag{1}$$

where $h_{k,j}$ is the percentage of AUM allocated to style segment k; $h_{i(k),j}$, the percentage of $h_{k,j}$ allocated to stock i in style segment k; $prc_{i(k)}$ and $shrown_{i(k),j}$, the price and shares owned of stock i respectively. In (1),

$$h_{k,j} = \frac{\sum_{i(k)\epsilon I(j,k)} (shrown_{i(k),j} \cdot prc_{i(k)})}{\sum_{k\epsilon K} \sum_{i(k)\epsilon I(j,k)} (shrown_{i(k),j} \cdot prc_{i(k)})} \tag{2}$$

[13]Specifically, industry-adjusted book-to-market is computed as $[ln(BTM_{i,t}^{i}) - ln(BTM_{t}^{j})]/\sigma[ln(BTM_{i,t}^{i}) - ln(BTM_{t}^{j})]$ where $BTM_{i,t}^{i}$ is the book-to-market ratio of stock i that belongs to industry j at June end of year t and BTM_{t}^{j} is the aggregate book value of stocks i in industry j divided by aggregate market value of stocks i in industry j at June end of year t.

$$h_{i(k),j} = \frac{(shrown_{i(k),j} \cdot prc_{i(k)})}{\sum_{i(k)\epsilon I(j,k)}(shrown_{i(k),j} \cdot prc_{i(k)})} \tag{3}$$

Similarly, let $\bar{h}_k$ denote the percentage of AUM aggregated across all funds allocated to style segment k, and $\bar{h}_{i(k)}$, the percentage of $\bar{h}_k$ allocated to stock i in style segment k.

$$\bar{h}_k = \frac{\sum_{j\epsilon J(k)}\sum_{i(k)\epsilon I(j,k)}(shrown_{i(k),j} \cdot prc_{i(k)})}{\sum_{k\epsilon K}\sum_{j\epsilon J(k)}\sum_{i(k)\epsilon I(j,k)}(shrown_{i(k),j} \cdot prc_{i(k)})} \tag{4}$$

$$\bar{h}_{i(k)} = \frac{\sum_{j\epsilon J(k)}(shrown_{i(k),j} \cdot prc_{i(k)})}{\sum_{j\epsilon J(k)}\sum_{i(k)\epsilon I(j,k)}(shrown_{i(k),j} \cdot prc_{i(k)})} \tag{5}$$

where $J(k)$ are the set of funds who own stocks in style segment k. From (1),

$$h_{i,j} = \left(h_{i(k),j} \pm \bar{h}_{i(k)}\right)\left(h_{k,j} \pm \bar{h}_k\right) \tag{6}$$

$$= \left(h_{i(k),j} - \bar{h}_{i(k)}\right)\left(h_{k,j} - \bar{h}_k\right) + \left(h_{i(k),j} - \bar{h}_{i(k)}\right)\bar{h}_k + \left(h_{k,j} - \bar{h}_k\right)\bar{h}_{i(k)} + \bar{h}_{i(k)}\bar{h}_k$$

$$\cong \left(h_{i(k),j} - \bar{h}_{i(k)}\right)\bar{h}_k + \left(h_{k,j} - \bar{h}_k\right)\bar{h}_{i(k)} + \bar{h}_{i(k)}\bar{h}_k \tag{7}$$

In (7), the first term represents the fund's portfolio return attributed to trading on firm-specific information, and the second term, to trading on systematic factors associated with size, book-to-market, and momentum stock characteristics. These two terms represent the principal components of variation in fund-stock holdings. The interaction term $\left(h_{i(k),j} - \bar{h}_{i(k)}\right)\left(h_{k,j} - \bar{h}_k\right)$, which represents a residual variation, is (approximately) 0 when stock selection and timing decisions are (largely) independent.

Active stock ownership associated with stock selection, w_{ij}, is the deviation of fund from peer group ownership of stock i in style segment k.

$$w_{ij} = h_{i(k),j} - \bar{h}_{i(k)} \tag{8}$$

Fund managers are more (less) optimistic than the overall market when the percentage of aggregate AUM in a style segment allocated to a stock in the style segment, $\bar{h}_{i(k)}$, is higher (lower) than the market capitalization of the stock relative to the aggregate market capitalization of stocks in a style segment. Deviations of fund from peer group ownership of a stock in a style segment will reflect differences in beliefs across fund managers.[14] For fund $j\epsilon J(k)$, the sum of active stock ownership across stocks i in style segment k, $\sum_{i(k)\epsilon I(j,k)} w_{ij} = \sum_{i(k)\epsilon I(j,k)} (h_{i(k),j} - \bar{h}_{i(k)}) = 0$. Moreover, for fund $j\epsilon J(k)$, deviations of fund from peer group ownership on every stock i in a style segment k will be zero only when fund managers are symmetrically informed.

B. Stock Investment Quality

When fund managers are unskilled or predominantly trade on sentiment, we expect active stock ownership to be uncorrelated with skill. The dispersion in active stock ownership from sentiment-based herding by unskilled fund managers which drive prices away from fundamental value predict lower future stock returns.[15] When skilled fund managers have private information and faced with the same information act similarly, we expect active stock ownership to be positively correlated with skill and co-movements in active stock ownership to reflect differences in private information about future stock return between skilled and unskilled fund managers.

We estimate the correlation of active stock ownership associated with information asymmetry and latent managerial skill associated with fund quality as:

$$IQ_i = \frac{\sum_{j\epsilon J(i,k)} (w_{ij}.\alpha_j)}{\sum_{j\epsilon J(i,k)} w_{ij}^2} = \sum_{j\epsilon J(i,k)} (\hat{w}_{ij}.\alpha_j) \tag{9}$$

In (4), $\hat{w}_{ij} = w_{ij} / \sum_{j\epsilon J(i,k)} w_{ij}^2$ is the active stock ownership of stock i in style segment k by fund j normalized by $\sum_{j\epsilon J(i,k)} w_{ij}^2$, the variance in the dispersion of active stock ownership across funds $j\epsilon J(i,k)$ associated with the degree of information asymmetry among fund managers. For high quality funds, stock IQ is higher (lower) when positive (negative) deviations in fund from peer group stock

[14]Alternatively, peer group ownership can be defined as the percentage of a fund's AUM in a style segment allocated to a stock in the style segment averaged across funds.

$$\bar{h}_{i(k)} = J(k)^{-1} \sum_{j\epsilon J(k)} \{(shrown_{i(k),j} \cdot prc_{i(k)}) / \sum_{i(k)\epsilon I(j,k)} (shrown_{i(k),j} \cdot prc_{i(k)})\}$$

Deviations of fund from peer group ownership will also sum to zero since $\sum_{i(k)\epsilon I(j,k)} \bar{h}_{i(k)} = 1$. But deviations will entwine differences in beliefs across managers on a stock in a style segment with differences in the size of AUM in a style segment across funds.

[15]Interpreting dispersion in analysts' forecasts as a proxy for differences in opinion, Diether, Malloy and Scherbina (2002) find that future returns are lower on stocks that exhibit higher dispersion in analysts' earnings forecasts. They find the dispersion effect to be most pronounced on stocks that performed poorly in the past year.

ownership are positively correlated with fund quality. For low quality funds, we expect deviations in fund from peer group stock ownership to be largely uncorrelated with fund quality and spurious correlations with fund quality will indicate noise trading.

> *Hypothesis 1a*: Stock IQ will strongly persist when active stock ownership is motivated by trades of privately informed skilled fund managers. Active stock ownership will be firmly and positively correlated with skill when stock IQ is high.
>
> *Hypothesis 1b*: Further, stock IQ will predict high future stock returns when the stock is widely held in common by skilled and privately informed fund managers.

To describe differences in fund manager skill that is consistent both across stocks and across quarters, we rank stocks by $IQ_i = \sum_{j\epsilon J(i,k)}(\widehat{w}_{ij}.\alpha_j)$. The percentile rank of stock i, p_i^s, is the fraction of all stocks where fund manager skill is less than or equal to the fund manager skill on stock i, and $(1 - p_i^s)$, the fraction of all stocks where fund manager skill is greater than that on stock i. We use an odds ratio, the relative percentile rank of stock i, $\theta_i^s = p_i^s/(1 - p_i^s)$, to proxy for the IQ of stocks.

C. Holdings Turnover and Conviction Quality

We also examine whether the selection skill of high performing fund managers is related to how frequently funds change their active stock ownership. Cremers and Pareek (2016) find that active share alone is not sufficient for fund managers to outperform. Only the most active and patiently managed funds outperform.[16] The conviction of fund manager beliefs on future stock returns is reflected in patience.

In current literature, fund turnover is proxied either by duration of holdings, reported fund

[16]See p. 295 in Cremers and Pareek (2016). In Pastor, Stambaugh, and Taylor (2017), fund turnover refers to frequent trading of stockholdings by funds rather than to the holding period of fund stockholdings or changes in active share in Cremers and Pareek (2016).

turnover ratio, or holding turnover. We consider each possible choice in turn. First, duration of stock ownership, which is the number of quarters a stock is held by a fund from ownership inception to the current quarter weighted by the percentage of shares outstanding each quarter (Cremers et al., 2016; and Lan, Moneta, and Wermers, 2019), has significant drawbacks. Fund age will bias holding horizons. On the same stock, inception dates will be earlier for mature funds compared to newly established funds. Second, changes in duration are capped and highly predictable. Holding horizon can at most increase by a quarter at a time and changes in stock ownership are slow to change from quarter to quarter. Third, reported fund turnover ratio cannot describe quarterly changes in individual holdings. In the spirit of Gaspar, Massa, and Matos (2005), we use style segment-adjusted holding turnover to proxy fund manager conviction and patience.

We characterize a fund manager's stock ownership turnover on stock i by Δw_{ij}, changes in the active stock ownership of fund j in stock i from four-quarter prior. If stock i is not held by fund j four quarters prior, $w_{ij}(q-4)$ takes value of zero.

$$\Delta w_{ij} = w_{ij}(q) - w_{ij}(q-4) \tag{10}$$

Active stock ownership turnover, $\Delta \widehat{w}_{ij}$, is fund stock ownership turnover normed by the dispersion in fund from peer group stock ownership turnover.

$$\Delta \widehat{w}_{ij} = \Delta w_{ij} / \sum_{j \epsilon J(i,k)} \Delta w_{ij}^2 \tag{11}$$

where $j \epsilon J(i,k)$ denotes the set of funds who trade stock i in style segment k. We identify the latent patience of fund managers on a stock, $\widehat{\pi}_i$, by the cross-product of active stock ownership turnover $\Delta \widehat{w}_{ij}$ and fund quality $\widehat{\alpha}_{j,t}$ summed across funds.

$$\widehat{\pi}_i = \sum_{j \epsilon J(i,k)} \Delta \widehat{w}_{ij} \cdot \widehat{\alpha}_j \tag{12}$$

Stocks exhibit marked impatience when a stock is actively traded by fund managers and skilled fund managers trade actively.[17] Again, to meaningfully describe differences in the latent patience of fund managers that is consistent both across stocks and across quarters, we rank stocks by patience $\hat{\pi}_i$. The percentile rank of stock i, p_i^{π}, is the fraction of all stocks in which fund managers trade less actively than fund managers who own stock i, and $(1 - p_i^{\pi})$, the fraction of all stocks in which fund managers trade more actively than fund managers who own stock i. The relative percentile rank of stock i is an odds ratio, $\theta_i^{\pi} = p_i^{\pi}/(1 - p_i^{\pi})$. We use relative percentile rank, θ_i^{π}, to proxy the trading activity and conviction quality of fund managers on a stock.

Hypothesis 2: High relative percentile ranks on active stock ownership turnover indicate impatience and lack of conviction. Future returns will be lower on stocks with high trading activity and low conviction quality.

D. Fund Quality

High performing fund managers are more likely to hold high quality stocks (Cohen et al., 2005). As in Berk et al. (2015), we instrument the latent quality of fund management by GVA, the product of gross (pre-expense return) alpha and TNA under management. We use the monthly seasonally adjusted CPI index (1982-1984=100) constructed by the U.S. Bureau of Labor Statistics (BLS) to adjust TNA under management for inflation.

For each fund, we estimate Fama and French (1993) and Carhart (1997) 4-factor model gross alphas from rolling twelve-month time series regressions of monthly gross excess returns on monthly excess market return $(r_{mt} - r_{ft})$, size (SMB_t), book-to-market (HML_t), and momentum (UMD_t) factors.

$$r_{jt} - r_{ft} = \alpha_j + \beta_i \cdot \left(r_{mt} - r_{ft}\right) + s_j \cdot SMB_t + h_j \cdot HML_t + u_j \cdot UMD_t + \varepsilon_{jt} \tag{13}$$

Monthly gross (pre-expense) fund return, r_{jt}, is the net monthly fund return plus one-twelfth of the

[17] Note that $E\left(\Delta\widetilde{w}_{ij} \cdot \tilde{\alpha}_{j,t}\right) = E\left(\Delta\widetilde{w}_{ij}\right)E\left(\tilde{\alpha}_{j,t}\right) + Cov\left(\Delta\widetilde{w}_{ij}, \tilde{\alpha}_{j,t}\right)$.

fund's annual expense ratio. From CRSP, the risk-free rate, r_{ft}, is the one-month Treasury bill yield at the beginning of month t. $(r_{mt} - r_{ft})$, SMB_t, HML_t, and UMD_t are the monthly market excess return, size, book-to-market, and momentum factors obtained from Ken French's website.

Monthly gross value-added, $\widehat{GVA}_{j,t}$, is the product of current month 4-factor alpha and prior month end TNA, as in Berk et al. (2015). To mitigate the volatile effects of transitory factors on long-term performance, gross value-added is time-averaged across current and prior months when the fund is in the sample, $T^{-1}\sum_{t=1}^{T}\hat{\alpha}_{j,t}$. Quarterly gross value-added is monthly gross value-added summed across months in the quarter.

When investors can detect skill, their allocation decisions determine fund size and managerial compensation. Net alpha is endogenously determined in equilibrium by competition among investors, and gross alpha, by the fees charged by funds. When managerial skill is in short supply and exhibit diminishing returns to scale, net alpha is driven to zero and managerial compensation is equal to gross value-added, the fund's gross excess return multiplied by total net assets under management. Gross alpha differentiates managers only when fees are such that all funds are the same size.

III. Data

A. Mutual Fund Sample

We select our sample of U.S. actively managed domestic equity mutual funds from the CRSP Survivor-Bias-Free Mutual Fund Database. Because mutual funds can have multiple share classes with the same underlying stock holdings, we use the database variable CRSP_CL_GRP to consolidate different share classes into a single fund as in Cao and Wermers (2018). Average total net assets (TNA) under management is TNA summed across underlying share classes each quarter, and monthly return is a TNA-weighted sum of underlying share class returns. As in Doshi, Elkamhi, and Simutin (2015), we use the database variable CRSP_OBJ_CD to identify domestic equity mutual funds and exclude sector funds, foreign funds, fixed income funds as well as mixed style funds. We use index identifier and fund names to exclude ETFs and ETNs, as well as index mutual funds. Akin to Kacpercyzk et al. (2008), we also exclude funds who, on average over our sample period, own fewer than 10 stocks or manage less than $5 million in TNA.

We obtain quarterly mutual fund holdings from the Thomson Reuter Mutual Fund Holdings database. We link actively managed domestic equity mutual funds in our sample with stock holdings data through MFLINKS. We exclude funds we could not match. For funds with missing reports in four or less quarters, we linearly interpolate their holdings using the latest holdings available before and after the missing reporting period.

To compute stock level variables, we link the merged fund stock holdings to the CRSP stock database to obtain daily and monthly returns, price, volume, shares outstanding and other variables. We focus on common stocks with share code 10 or 11 that trade on NYSE, NASDAQ, or AMEX. We adjust stock trading volumes in NASDAQ broker-dealer markets reported in CRSP by one-half following French (2008). We also link stock holdings to the S&P Compustat database to obtain book-to-market ratios. To moderate the influence of outliers on our results, we only keep stocks held by at least five mutual funds in the quarter and eliminate stocks with share prices below $5.

Data availability constraints on CRSP_CL_GRP and WFICN in MFLINKS restrict our sample period to start in 2000 and end in 2017. A set of 2,224 unique mutual funds who collectively own 7,447 unique stocks meet our screening criteria. Over our sample period, the number of mutual funds rose from 897 to 1,331, and number of stocks owned fell from 3,422 to 2,927.

B. Characteristics of Style Segments

The characteristics of fund-stock ownership across the 125 style segments over our 72-quarter (2000-2017) sample period are summarized in Table I. Column 1 denotes size quintiles from small to large, and Column 2, book-to-market quintiles from low to high. Top row denotes momentum quintiles from low to high.

< Insert Table I here. >

Table I Panel A reports the median number of CRSP stocks in our sample that fall into each style segment, and in parentheses, the median stock ownership across funds in a style segment. Stock ownership is the number of stocks a fund owns expressed as a percentage of all stocks in a style segment. The median number and median stock ownership increases across momentum quintiles but only in the smallest size quintile. Two trends are apparent from the $Size\_BTM$ column, which reports median number and median stock ownership averaged across momentum quintiles on stocks sorted into size and book-to-market quintiles. The median number of stocks decreases across size quintiles, ranging from 46 to 56 in the smallest size quintile to 13 in the largest size quintile. At the same time, stock ownership increases across size quintiles, ranging from 3.4% to 3.8% in the smallest size quintile to 17.8% in the largest size quintile.

Table I Panel B reports the median number of funds who own stocks in the style segment, and in parentheses, the median fund ownership across stocks in a style segment. For each stock in a style segment, fund ownership is the number of funds who own the stock expressed as a percentage of the total number of funds in the style segment. With a few exceptions, the median number of funds increase across momentum quintiles. Median fund ownership also increases across momentum quintiles except in the smallest size quintile. In the smallest size quintile, fund ownership averaged across book-to-market quintiles is 8.3% on high momentum stocks compared to 9.5% on low momentum stocks.

Two trends are evident from the $Size\_BTM$ column, which reports median number of funds and median fund ownership averaged across momentum quintiles on stocks sorted by size and book-to-market quintiles. The median number of funds and median fund ownership do not vary notably across book-to-market quintiles. The median number of funds and median fund ownership, however, increase across size quintiles. In number, ranging from 172 to 232 in the smallest size quintile, and in the largest size quintile, from 574 to 692. In ownership, ranging from 8.3% to 9.1% in the smallest size quintile, and in the largest size quintile, from 19.8% to 21.0%.

It is apparent the small number of stocks in the largest size quintile attract the largest number of funds, and fund ownership is also highest. Stocks in the largest size quintile have the deepest breadth and ownership. Breadth and ownership are higher as firm market capitalization grows bigger.

< Insert Figure 1. >

Figure 1 graphs the distributions of stock and fund ownerships averaged across momentum quintiles on stocks sorted first by size (x_1), and secondly, by book-to-market (x_2) quintiles reported in the average columns in Table 1. The coordinate vectors $(x_1 x_2) = (k_1 k_2)$ on the $Size\_BTM$ axis denote the $k_1 = 1, \dots, 5$ size quintiles and $k_2 = 1, \dots, 5$ book-to-market quintiles. Symbols ○ and ◊ denote the variables of interest whose values are plotted along the left and right scales on the vertical axis.

A risk diversification motive is discernable. From Figure 1 Panel A, as market capitalization increases, the median number of stocks funds own declines from a high of 56 to a low of 13, and median stock ownership rises from a low of 3.4% to a high of 17.8%. Additionally, from Figure 1 Panel B, as market capitalization increases, the percentage of funds who own a stock, increases from a low of 8.3% to a high of 21%. The effect of book-to-market on stock and fund ownership is generally weak. Funds are more likely to own large cap stocks, and fund ownership is more concentrated in large cap stocks. Overall, stock and fund ownerships are higher on large cap stocks.

In summary, Table I shows that the average number of stocks that funds own as a percentage of all stocks in a style segment (stock ownership) and the average number of funds who own a stock as a percentage of the total number of funds in the stock's style segment (fund ownership) are lowest in the smallest size quintile and highest in the largest size quintile.

We also examine average fund quality across style segments. Results are reported in Appendix

Tables II, and Online Appendix Figures I and II. Our findings substantiate Berk and Binsbergen (2015). In competitive markets for investible funds by investors who can detect managerial skill, net alpha is endogenously determined by fees. Gross value-added is a better proxy of fund quality. With diminishing returns to scale, gross alphas initially increase but eventually decrease with TNA. GVA will be a strictly concave function of TNA under management (see Zhu, 2018: Figure 2). In the remainder of the paper, we use GVA to instrument fund quality. We examine management fees as an alternative proxy for fund quality. A more detailed discussion can be found at the end of the paper.

C. Summary Statistics

Table II reports summary statistics on variables used in our analysis. To mitigate the effect of outliers, all variables are winsorized at the top and bottom 1%. Variables are defined in the table II heading and summarized as well in Appendix Table I.

< Insert Table II here. >

As an alternative proxy for fund quality, we use management fees (Berk et al., 2015), which is estimated each month as the product of fund TNA at the end of the prior month and 1/12 of the annual management fee ratio as a percentage of fund TNA reported by CRSP, and time-averaged from the start of the sample period. Monthly management fees are summed over three months in a quarter to compute quarterly management fees. Both GVA and management fees are measured in dollars. Higher performing funds can extract higher management fees from investors. In a competitive market where investors are able to identify higher quality funds, fund return premia will be driven to zero. Management fees will equal GVA. Note that because we proxy selection and conviction qualities by relative percentile ranks, the selection quality and conviction measures have identical distributions across the whole sample as shown in Table II. The identical distributions make coefficients in subsequent regressions comparable and easy to interpret.

To fairly judge the contribution of IQ to forecast future stock returns, we take other documented empirical predictors into account. As in Chen, Hong, and Stein (2002), we use delta breadth and delta mutual fund ownership to proxy for differences in opinion and short-sale constraints. Low breadth and low institutional ownership signal short-sale constraints are tightly binding, and prices are high relative to fundamentals. Increases in delta breadth and institutional ownership should forecast higher returns. Following Jiang and Sun (2014), we compute breadth as $ln(1+N)$ where N denotes the number of funds who own the stock, and mutual fund ownership as the fraction of total shares outstanding owned by actively managed mutual funds. Quarterly changes in breadth and active mutual fund ownership are computed as the change in breadth and ownership from the prior quarter. The mean (median) breadth of 3.606 (3.689) suggests an average (median) of 36 (39) active mutual funds own a stock in our sample. Our mean (median) breadth is higher than 25 (11) reported in Jiang and Sun (2014) because of the significant growth in the number and size of funds in our more recent sample period 2000 to 2017 in contrast to their sample period 1984 to 2008.

Jiang and Sun (2014) and Jiang et al. (2020) find that disagreement and consensus in opinion among fund managers predict future stock returns. Akin to the dispersion in analysts' opinions in Diether, Malloy, and Scherbina (2002), we compute a dispersion index of active holding as the standard deviation of active holding divided by absolute value of mean in active holding, which has a sample mean (median) of 0.137 (0.139).

IV. Active Management and Future Stock Returns

A. Persistence in Stock Investment Quality

If stock IQ imbeds the co-movements in active stock ownership from trading on private information by skilled fund managers rather than from sentiment-based trading by unskilled fund managers, we expect stock IQ to exhibit persistence. We examine persistence in two approaches.

First, we sort all stocks into deciles by their IQ at the end of each quarter and compute the average IQ across all stocks by decile in lead two, four, eight, and twelve quarters. Spreads in average IQ between the top and bottom deciles and their t-statistics are reported in Table IV Panel A. Decile spreads decrease over the next twelve quarters but remain highly significant.

< Insert Table III here. >

Second, we examine the persistency of stock IQ through a transition table of quarterly changes in stock IQ. At the end of the prior and current quarters, we sort stocks into quintiles by IQ and compute the fraction of stocks that move from quintile i in quarter $t-1$ to quintile j in quarter t. The transition matrix reported in Panel B, which has a dominant diagonal, converges. The likelihood that stocks in the bottom quintile remain in the bottom quintile is 75.74%, and the likelihood that stocks in the top quintile remain in the top quintile is 71.86%. Stock IQ is strongly persistent.

Table III Panel C reports summary statistics on active stock ownership, GVA, market capitalization, book-to-market, and Pearson rank correlations between fund active stock ownership and GVAs at the 5th through 95th percentiles. Negative active stock ownership and negative correlations between active stock ownership and GVAs indicate that skillful fund managers underinvest relative to their peer group on stocks at the lowest quintile of IQ. Active stock ownership and the correlations between active stock ownership and GVAs become increasingly more positive on stocks ranked higher on IQ. These corroborating results show that high IQ of stocks are in the hands of more skillful and better-informed managers. The focus of stock selection at upper percentile ranks of IQ appears to be on growth rather than value stocks, and in middle percentile ranks of IQ, on small rather than large cap stocks.

B. Comparison with Alternative Fund Holdings-Based Stock Return Predictors

As in Wermers, Yao, and Zhao (2012), we use Fama-Macbeth (1973) regressions, corrected for correlated errors using a Newey-West estimator with one quarter lag, to examine the forecast return power of stock IQ in a horserace against four widely-cited empirically documented stock return predictors: herding by unskilled fund managers, adjusted 'dumb' money flow of investor-sentiment driven trading, as well as delta breadth and delta ownership in mutual fund holdings that reflect short-sale constraints. We compute herding in fund holdings following Lakonishok, Shleifer, and Vishny (1992, eq. 1) using an adjustment factor in Jones, Lee, and Weis (1999). Delta breadth and delta mutual fund ownerships are estimated following Chen, Hong, and Stein (2002). Adjusted money flow is computed following Frazzini and Lamont (2008, eq. 8) with one-month horizon and summed across three months in a quarter.

< Insert Table IV here. >

Regression results are reported in Table IV. Controlling for competing stock return predictors, the forecast power of stock IQ remains significant. The forecast return persistence of stock IQ substantiates trading by skilled fund managers on private information. The coefficients on stock IQ are significantly positive over lead four quarters. The significantly positive coefficients on delta breadth confirms Chen, Hong, and Stein (2002), and the relatively weak feedback effects on stock returns from mutual fund managers' buy and sell herding measures is consistent with Lakonishok, Shleifer, and Vishny (1992).

C. Portfolio Buy-and-Hold Returns: One-way Sorts

At the end of each quarter, we assign stocks into IQ sorted quintile portfolios using the stock's relative percentile rank, θ_i. We compute value-weight monthly and average quarterly buy-and-hold returns on each IQ quintile portfolio in the lead month and quarters following quarter-end portfolio formation and link quintile portfolio returns to form a time-series. Monthly and average quarterly quintile portfolio returns averaged across our sample period on the one-way sorts of stocks by IQ are reported in Table V.

< Insert Table V here. >

As evident in Table V, future stock returns increase with higher IQ. A value-weight portfolio of high IQ stocks outperforms a value-weight portfolio of low IQ stocks. In the subsequent quarter, portfolio returns are consistently negative in the bottom quintile, and positive in the top two quintiles of IQ. The forecast average quarterly excess market return is 1.533%, DGTW return is 1.723%, and 4-factor alpha is 1.537%. Because active stock ownership on IQ are based on style segments, positive spreads on high-low quintile portfolios of stocks by IQ cannot be attributed to size, book to market, or momentum stock characteristics. Forecast returns are robust to alternative adjustments for risk and strongly persist through a twelve-month period.

D. Portfolio Buy-and-Hold Returns: Two-way Sorts

Table VI reports average quarterly portfolio returns on two-way sorts of stocks, first into terciles by NYSE market capitalization, and second, into IQ quintile portfolios by a stock's relative percentile rank, θ_i. To prevent microcaps from driving portfolio outperformance, Hou, Xue, and Zhang (2020) stress the importance of using NYSE breakpoints for market capitalization and value-weighted portfolio returns (Fama and French, 1993). Portfolios sorted by NYSE breakpoints exhibit more consistency over time. At the end of each quarter, we report average value-weight buy-and-hold-returns on $3 \times 5 = 15$ portfolios following portfolio formation, in lead one to four quarters. The quintile portfolio returns are linked to form a time-series.

< Insert Table VI here. >

A long high-short low portfolio trading strategy generate statistically and economically significant excess quarterly returns across all market capitalizations on IQ. As apparent from Table VI, across all terciles of market capitalization, portfolio returns are consistently negative in the bottom quintile of IQ, and predominantly positive in the top two quintiles of IQ. Value-weight portfolios of high IQ stocks outperform low IQ stocks across all terciles of market capitalization. Forecast returns are higher on small cap stocks and lower on mid- and large-cap stocks for excess market return and DGTW return. On small cap stocks, average quarterly excess market return is 1.825%, DGTW return is 1.907%, and 4-factor alpha is 0.670%, in lead one quarter. In comparison, on large cap stocks, average quarterly excess market return, DGTW return, and 4-factor alpha are 1.646%, 1.607% and 1.241% in lead one quarter.

The information advantage of skillful managers is more muted in large cap stocks which attract more attention, are more closely scrutinized, and more transparent. In subsequent multivariate regressions we corroborate the finding that forecast returns are greater on small-cap stocks and smaller on large-cap stocks with higher IQ.

V. Forecast Return Regressions

In this section, we estimate two-way style and quarter fixed effects regressions of lead quarter stock returns on IQ, controlling for delta breadth, delta active mutual fund ownership, dispersion in fund active holdings, natural logs of market cap and book-to-market, prior 12-month return, CRSP turnover, idiosyncratic volatility, and market beta. Errors are clustered by style and quarter. We add a squared IQ term to account for possible diminishing returns on IQ. Quarter returns in regression tables are expressed in percent. We winsorize all variables at the 1st and 99th percentiles. For ease of interpretation, all regressors are normalized by their standard deviations across the sample period, and control variables are demeaned. Investment style and time fixed effects are added to control for unobservable time-invariant style segment factors, as well as quarterly unobserved common factors. Results are reported in Table VII.

< Insert Table VII here. >

Forecast quarterly returns are significantly and economically greater on stocks ranked higher on IQ. In Table VII Panel A, estimated coefficients on IQ relate forecast quarterly returns to standardized units of IQ. Forecast quarterly return is the product of estimated coefficient and relative percentile rank scaled by the sample standard deviation of relative percentile ranks. We can compute relative percentile rank θ_i from percentile rank $p_i = \theta_i/(1+\theta_i)$. The forecast quarterly returns on a stock at percentile ranks p_i ranging from 0.05 to 0.95 are shown in Table VII Panel B.

A standard deviation increase in IQ will raise average forecast quarterly excess market return by 1.360%. In lead one quarter, average forecast quarterly excess market return at the mean IQ is 0.487% (=1.360*4.572/12.760), where 4.572 and 12.760 are the mean and standard deviation of IQ reported in Table III. To put this in perspective, a 5th to 95th percentile change in IQ increases average forecast quarterly excess market returns by 2.019% (=1.360*(19.0-0.053)/12.760), while a 90th to 95th percentile change, by 1.066% (=1.360*(19.0-9.0)/12.760). A percentile change in IQ has a greater impact on forecast returns at higher percentiles.

Results in Table VII Panel B corroborate Table V. A portfolio of stocks in the highest quintile by relative IQ percentile rank outperforms a portfolio of stocks in the lowest quintile rank by 1.533%, as reported to Table V, which is close to but slightly smaller than the 5^{th} to 95^{th} percentile spread. From Panel B of Table VII, in lead one quarter, the forecast excess market return spread on a long-short portfolio of stocks in the 95^{th} and 5^{th} percentile is 2.019%. The spreads on DGTW return and 4-factor alpha are 1.979% and 1.329%.

The magnitude of forecast quarterly returns from IQ decline each quarter but strongly persist through four quarters. In the fourth quarter, average quarterly excess market return, DGTW return, and 4-factor alpha are 25.15% (=(1.360-1.018)/1.360), 28.89%, and 14.41% lower compared to first quarter returns. The information advantage of fund managers from selection skill decay slowly.

We control for delta breadth and delta ownership (Chen et al., 2002). Estimated delta breadth coefficients, which are statistically significant at the 1% level in lead one and two quarters, are comparable to those in Table 6 Panel A of Chen et al. (2002). A standard deviation increase in delta breath forecasts a higher lead one quarter excess market return of 0.533%, DGTW return of 0.497%, and 4-factor alpha of 0.312%. On returns unadjusted for market return or risk, Chen et al. (2002) document a standard deviation increase in delta breadth forecasts a higher return of 0.546% (=1.187*0.46%) where 0.46% is the standard deviation of delta breadth reported in Table 1 of Chen et al. (2002). Note however that estimated coefficients on delta breadth lose statistical significance after the second quarter. Finding is consistent with McLean and Pontiff (2016), who show that as investors learn about and trade on empirically documented predictors, performance decays post-publication. As in Chen et al. (2002), we also find that controlling for delta breadth, delta ownership is insignificant.

Larger standard deviations in fund from peer group stock ownership forecast lower future stock returns. Estimated coefficients on dispersion index in active stock ownership are negative but statistically insignificant in lead quarters. Finding of lower future stock returns is consistent with Diether et al. (2002) who show that higher dispersions in analysts' earnings forecasts indicate more uncertainty about fundamental value, and Miller (1997), that constraints on short-sales cause stock prices to be high relative to intrinsic value.

We use CRSP turnover to account for a possible performance-turnover relation (Brennan, Chordia and Subrahmanyam, 1998; and Pastor, Stambaugh, and Taylor, 2020). The estimated coefficients on CRSP turnover are negative and significant in lead two, three, and four quarters consistent with a negative effect of trading volume on stock returns documented in Brennan et al. (1998). CRSP turnover is also a proxy of stock liquidity, and the negative coefficients on CRSP turnover are consistent with that more liquid stocks require lower returns. In lead second quarter, a standard deviation increase in CRSP turnover predicts lower excess market return of 0.443%, DGTW return of 0.263%, and 4-factor alpha of 0.441%. Excess returns in lead three and four quarters from an increase in CRSP turnover are slightly more negative.

Negative coefficients on squared relative percentile IQ rank, ${\theta_i^s}^2$, indicate diminishing returns to investment quality. Marginal reductions in forecast quarterly return are shown in Table V Panel B at the 5th through 95th percentile ranks. On average, the impact is much smaller relative to the main effect of IQ on forecast returns.

Lastly, Table VII Panel C reports Fama-Macbeth (Fama and Macbeth, 1973) regressions over our sample period. Newey-West (1987) standard errors are estimated assuming a one-quarter lag in serially correlated errors. The coefficients on IQ closely resemble those in two-way style and quarter fixed effects regressions in Table VII Panel A. Overall, results are robust to alternative model estimation methods.

B. Alternative Proxy for Fund quality

Berk et al. (2004, 2015) argue that in a market where skill is in short supply, more skilled fund managers can choose the fees they charge investors. When investors can detect skill, net alpha is driven to zero in equilibrium by competition among investors. In equilibrium, fund gross alpha will equal fees charged.

We use management fees as an alternative proxy for fund quality. Each month, management fee is the product of a fund's TNA in the prior month end and one-twelfth of its annual (fiscal year) management fee, which is subsequently time-averaged from the start of the sample period to the current month. We sum up monthly management fees over three months in the quarter to obtain quarterly management fees. Each quarter, stock IQ based on management fees is estimated as the cross-product of active holding and management fees summed across funds. As previously, we rank stock IQ based on management fees and compute the odds ratio.

< Insert Table VIII here. >

Two-way fixed effects regression results reported in Table VIII mirror Table VII. Forecast quarterly returns using management fee-based stock IQ strongly persist through four quarters. In lead one quarter, a standard deviation increase in management fee-based stock IQ increases average forecasted quarterly excess market return, DGTW return, and 4-factor alpha on stocks by 1.328%, 1.364%, and 0.881%. Similar to Table VII Panel B, a long-short portfolio of stocks in the top 95th and bottom 5th percentile generates a quarterly excess market return of 1.972% (=1.328*(19.000-

0.053)/12.760), where 12.760 is the standard deviation of IQ. Findings corroborate Berk et al. (2004, 2015) that fund managers can extract rent for skill through higher management fees. Slightly larger forecast quarterly returns in the third and fourth quarters on stocks ranked by IQ proxied by management fees suggest more skillful fund managers are able to charge more for their services, and as Zhu (2018) points out, is possible when investors can only discover managerial skill over time.

As a robustness test, we use industry concentration to proxy for fund quality. Kacperczyk et al. (2005) document significant diseconomies of scope. Skilled fund managers can exploit their information advantage and achieve superior performance by holding more concentrated industry portfolios. Each quarter, industry concentration is computed as the squared differences between industry weights of funds, $w_{j,\mathrm{I}}$, and aggregate industry weights, w_{I}, summed across ten broadly defined industries (Fama and French, 1997).

$$Industry\ concentration_{j,\mathrm{I}} = \sum_{\mathrm{I}=1}^{10}(w_{j,\mathrm{I}} - w_{\mathrm{I}})^2 \tag{12}$$

Aggregate industry weight is the dollar value invested in an industry as a percent of the total dollar value across all industries, aggregated over all sample funds. IQ is estimated in the same spirit as the sum product of fund quality and active holdings, with fund quality constructed by industry concentration time-averaged across current and prior quarters. Two-way stock and quarter fixed effects regression results, which are reported in Appendix Table III, closely resemble those in Table VII and Table VIII. Overall, the predictive return content of IQ is robust to management fee-based or industry concentration-based proxies of fund quality.

C. Market Capitalization and Investment Quality

Table IX examines the impact on forecast stock returns from IQ on stocks categorized by market capitalization. Using breakpoints on NYSE stocks, we sort stocks into terciles by market capitalization. The interaction between investment quality and market cap dummy is $\theta_i \cdot D_{i\kappa}$, where $D_{i,1} = 1$ if stock i is small cap and 0 otherwise, $D_{i,2} = 1$ if stock i is midcap and 0 otherwise, and $D_{i,3} = 1$ if stock i is large cap and 0 otherwise. Results of two-way stock and quarter fixed effects regressions of forecast stock returns on IQ among small cap, midcap, and large cap stocks are reported in Table IX.

< Insert Table IX here. >

As evident in Table IX Panel A, future stock returns are higher on stocks with better IQ regardless of market capitalization. Finding corroborates Table VI. IQ is, however, more important on small cap stocks. In lead one quarter, a standard deviation increase in IQ, increases average forecast quarterly excess market return on small cap stocks by 1.569%, compared to 1.306 % on midcap stocks, and 1.114 % on large cap stocks. Estimated forecast quarterly returns on the IQ of midcap and large cap stocks are on average 83.24% and 71.00% of those on small cap stocks. The same pattern is true for DGTW return and 4-factor alpha.

Further, forecast quarterly returns from IQ persist over four quarters, and small cap stocks benefit most from IQ over longer holding horizons. The quarterly returns with each additional quarter decline over small cap, midcap stocks, and large cap stocks, with the decay in large cap stocks much more attenuated. On small cap stocks, quarterly excess market return in the fourth quarter is lower than in the first quarter by 27.66% (=(1.569-1.135)/1.569), and on midcap stocks, lower by 30.25%. In comparison, quarterly excess market return in the fourth quarter is lowered by only 12.84% on large cap stocks. DGTW return and 4-factor exhibit the same pattern.

The forecast quarterly returns on small, mid, and large-cap stocks at percentile ranks p_i ranging from 5^{th} to 95^{th} are shown in Table IX Panel B. A portfolio of stocks ranked higher in IQ outperforms a portfolio of stocks ranked lower in IQ across all terciles of market capitalization. In lead one quarter, excess quarterly market return on a high-low IQ portfolio of small cap stocks is 2.330%, compared to 1.939% on midcap stocks, and 1.654 % on large cap stocks. Return spreads are similar on DGTW return and 4-factor alpha and strongly persist through the ensuing year.

As a robustness test, we use absolute forecast errors (AFE) to proxy for the value of private firm-specific information. More opaque and less transparent firms are more likely to have higher analyst forecast errors. Skilled fund managers can exploit their abilities to acquire private information on these firms and achieve outperformance by holding portfolios that are more concentrated in stocks with higher AFEs. Absolute forecast errors are estimated as the absolute differences of mean analyst forecasts for the period from realized earnings scaled by realized earnings. We separate firms with negative earnings in the quarter from firms with non-negative earnings because negative earnings are less likely to persist.

In two-way stock and quarter fixed effects regression results reported in Online Appendix Table I, the interaction terms between IQ and AFE tercile dummies as well as between IQ and a dummy for negative quarterly earnings summarize the effect of AFE. Results confirm that more opaque and less transparent firms with higher AFEs forecast higher future stock returns. Overall, the predictive return content of IQ is robust to market capitalization or AFE-based proxies for the value of private information.

D. Investment Quality and Conviction Quality

High fund active holdings turnover indicates less patience and low conviction quality. When patience and conviction of fund managers in a stock is important, we expect future returns to be lower on stocks with higher active stock ownership turnover. Results of style and quarter fixed effects regressions of forecast stock returns on IQ and trading activity are reported in Table X. All regressors are normalized by their standard deviations over the sample period and control variables are all demeaned.

< Insert Table X here. >

There are three panels in Table X. Panel A examines forecast quarterly returns from IQ and conviction quality, and Panel B, the interacted effect of IQ and conviction quality on forecast quarterly returns. Summary statistics on actively traded stocks in the 5th to 95th percentiles in conviction quality are reported in Panel C. For Panel B, Hi_IQ and Lo_IQ dummies denote stocks with above and below median IQ in the quarter. The interactions of Hi_IQ and Lo_IQ with conviction quality are used as regressors in Panel B model specifications.

Estimated coefficients on IQ in Table X Panel A are identical to coefficient estimates in Table VII Panel A. Accounting for the patience and conviction of fund managers does not diminish forecast quarterly returns from IQ. Combined with results in Panel B, on stocks where the trading activity of fund managers is high, however, forecast quarterly returns fall significantly. Further, forecast quarterly returns continue to decline each quarter. Compared to the first quarter, a standard deviation increase in trading activity will lead to a continuous fall in quarterly excess market return that returns in the fourth quarter are 53.92% (=0.227/0.421) of the first quarter, on above median IQ stocks. Results are similar for DGTW return and 4-factor alpha.

Active trading reduces forecast quarterly returns. The decline in forecast quarterly returns from trading activity shows the greatest profit from market mispricing occurs over longer holding periods. In Table X Panel B, forecast returns are higher when skillful fund managers are also patient investors. In lead one quarter, a standard deviation increase in conviction quality lowers average forecasted quarterly excess market return on a portfolio of above median IQ stocks by -0.421%. DGTW return and 4-factor alpha exhibit the same pattern, although the negative impact on 4-factor alpha is smaller in magnitude.

Patience is, however, not a substitute for skill. Forecast returns do not significantly increase when less skilled fund managers are also patient investors. Coefficients on the interactions of Lo_IQ and conviction quality are insignificant.

Lastly, Table X Panel C shows that for 80% of stocks, active turnover is negatively correlated with GVA. Correlation becomes more negative as active turnover increases from the 10^{th} percentile to the 50^{th} percentile and turns positive between the 80^{th} and 90^{th} percentiles of active turnover. Trading activity is predominantly by less skilled fund managers, and highest in mid cap-value stocks where skilled fund managers contribute to trading activity.

Our results conform to the Cremers et al. (2009) thesis that turnover by unskilled fund managers does not add value. Only funds with high active stock ownership and low active stock ownership turnover outperform their benchmarks. Forecast returns are the highest on high active ownership of stocks by skilled fund managers who are patient investors with long holding horizons. High turnover make unskilled fund managers look busy and creates buy-sell pressure that drive stock prices away from fundamental values and forecast lower future returns (Miller, 1997).

E. Investment Quality and Conviction Quality of Stocks by Market Capitalization

Table XI examines the impact of conviction quality on future returns of stocks categorized by market capitalization. Similar to Table IX, we sort stocks into terciles by market capitalization using breakpoints on NYSE stocks. Conviction quality is the relative percentile rank on active turnover. We construct interaction terms $\theta_i^{\pi} \cdot D_{i\kappa}$, where $D_{i,1} = 1$ if stock i is small cap and 0 otherwise, $D_{i,2} = 1$ if stock i is midcap and 0 otherwise, and $D_{i,3} = 1$ if stock i is large cap and 0 otherwise. Style and quarter fixed effects regressions of forecast returns on the IQ and conviction quality of small cap, midcap, and large stocks are reported in Table XI.

< Insert Table XI here. >

As expected, results confirm prior Tables IX and X. Future stock returns are higher on stocks with better IQ regardless of market capitalization. Estimated coefficients on IQ by market capitalization closely resemble those in Table IX. The information advantage of skilled fund managers is greater on small cap stocks and more muted in large and mid-cap stocks. Large cap stocks benefit least from IQ in short run and mid cap stocks benefit least in long run. As in Table X, accounting for the patience and conviction of fund managers does not diminish forecast returns from IQ.

High active stock ownership turnover and low patience adversely impact forecast returns on large cap stocks. In the first quarter, forecast quarterly returns on large cap stocks of fund managers who trade actively, are relatively lower by 0.720% on excess market return, 0.790% on DGTW return, and 0.745% on 4-factor alpha. Further, low conviction quality forecast lower returns that persist and increase in magnitude through the ensuing four quarters. Compared to the first quarter, average forecast quarterly excess market return, DGTW return, and 4-factor alpha on actively traded large cap stocks are relatively lower in the fourth quarter by 0.287% (=(-1.007%)-(-0.720%)), 0.034%, and 0.068%. Forecast returns are significantly lower in the fourth quarter, but insignificantly lower in the first quarter on midcap stock holdings for fund managers who trade frequently. For small cap stocks, forecast returns are significantly lower in most of the model specifications.

F. Investment and Conviction Quality by Fund Quality

Appendix Table II Panel C shows that fund quality falls approximately into terciles. Funds in the top tercile of fund quality have the highest (positive) GVA, and funds in the bottom tercile have the lowest (negative) GVA. If stock investment and conviction quality represent trading and turnover by skilled fund managers on private information, we should expect the forecast return power of investment and conviction quality to come mainly from funds in the top tercile of fund quality.

To test this thesis, we sort funds into terciles by GVA at the end of each quarter. We construct IQ as the sum products of active stock ownership and fund GVA across funds in each tercile. We compute active stock ownership turnover similarly by GVA tercile each quarter. Two-way fixed effects regressions of future stock returns for each GVA tercile are reported in Table XII.

< Insert Table XII here. >

Table XII confirms the forecast return power of investment and conviction quality come mainly from the top tercile of funds by GVA. In the top tercile of funds by GVA, IQ predicts significantly positive future stock excess market and DGTW returns as well as 4-factor alphas, which persist over four quarters. In the middle tercile, IQ predicts smaller and occasionally significantly positive future excess market and DGTW stock returns, and positive but insignificant 4-factor alpha. In the bottom tercile, IQ predicts insignificantly negative future excess market and DGTW stock returns, but significantly positive 4-factor alphas, which are smaller in magnitude than in the top GVA tercile. The results are not surprising. Funds are overfunded in the bottom tercile. As shown in Appendix Table II Panel C, net alphas do not correlate with GVA when funds are overfunded.

Coefficients on the conviction quality exhibit similar patterns. In the top tercile, low conviction quality-high $Turn$ predicts significantly negative future excess market and DGTW returns as well as 4-factor alphas. In the middle and bottom terciles, conviction quality predicts insignificant future stock returns excess market and DGTW returns as well as 4-factor alphas.

VI. Earnings Announcements

Lastly, we examine the private information of skilled fund managers, impounded in stock IQ, that is made public in earnings announcements. Earnings announcement dates are obtained from I/B/E/S database. We estimate CAR1 as the cumulative abnormal return over the three-day window [-1, 1] around the earnings announcement date, and CAR2 as the cumulative abnormal return over the window [3,60] from the 3rd day to the earlier of the day prior to the earnings announcement date in the subsequent quarter or 60th day post earnings announcement date. Abnormal daily returns are computed as daily returns in excess of returns on a 2×3 benchmark portfolio of stocks sorted by size (ME) and book-to-market equity (BE/ME) to which the stock belongs.[18] Two-way stock and quarter fixed effects regressions of CAR1 and CAR2 on stock IQ are reported in Table XIII. Errors are clustered by stock and quarter. Explanatory and control variables are normalized by their standard deviations over the sample period.

< Insert Table XIII here. >

[18]At June end of each year t, stocks are sorted into 2×3 benchmark portfolios by size (ME) and book-to-market equity (BE/ME). Median ME on NYSE stocks and 30th and 70th percentiles of BE/ME on NYSE stocks, computed as book equity in the last fiscal year end in $t-1$ divided by ME in December of $t-1$, are used as breakpoints.

The information content of earnings announcements in average CAR1 increases by 0.198% on a standard deviation rise in IQ. At the mean IQ, average forecast CAR1 in lead one quarter is 0.071% (=0.198*4.572/12.760), where 4.572 and 12.760 are the mean and standard deviation of IQ reported in Table III. To put this in perspective, forecast CAR1 is higher by 0.254% (=0.198*(19.0-0.053)/12.760) on a 5th to 95th percentile rise in IQ, and higher by 0.155% (=0.198*(19.0-9.0)/12.760) on a 90th to 95th percentile rise in IQ. Impact on forecast CAR1 is greater at higher percentiles of IQ.

< Insert Figure II here. >

CAR2 captures the private information in stock IQ made public in stock prices in the post-earnings periods following the three-day window around earnings announcements. The CAR for the top and bottom quintiles are graphed in Figure II. At the end of each quarter, we sort stocks into quintiles using stock IQ, and link each quintile over the sample period to form a time-series. We estimate the average CAR for each quintile in the lead one quarter post-earnings announcement period and compute the spread between the top and bottom quintiles.

The post-earnings announcement drift evident in CAR2 is consistent with a slow diffusion of fundamental information. When trading order imbalances create price pressure and signal informed trading when noise trading is low, informed traders will trade in such a way that their private information is incorporated into prices gradually (Kyle, 1985). The incidence of odd-lot trades in equity markets (O'Hara, Yao, and Ye, 2014) and order splitting (Bernhardt and Hughson, 2015) supports the slow diffusion of information.

Finally, Table XIII shows that post-earnings announcement CAR2 is greater on high IQ stocks as shown in Figure II. High IQ stocks embed more private information that is incorporated into prices around earnings announcement and post-earnings announcement periods. A standard deviation rise in IQ raises CAR2 in lead one quarter by 0.502%. IQ strongly predicts CAR2 up to lead four quarters.

VII. Conclusion

We show stock IQ derived from publicly available information on fund holdings and fund quality can be used to make more profitable investment decisions. Sorting stocks into style segments, deviations in fund from peer group stock ownership weighted by fund quality signals a stock's IQ. Managers of high performing funds, who are more skilled and better informed, make similar decisions when they act on the same information. Stocks ranked high on IQ generate significant positive excess market and risk-adjusted returns that persist through the ensuing year. The positive return-IQ relationship is robust to whether fund quality is proxied by fund GVA, management fees or industry concentration. Moreover, we show private information impounded in stock IQ is made public in earnings announcements.

In contrast, active stock ownership turnover and low patience predict lower future stock returns which also persist through the ensuing year. Future returns on high IQ stocks are adversely affected when fund managers lack patience and conviction of beliefs.

The forecast return power of investment and conviction quality come mainly from funds in the top tercile of fund quality, where stock selection ability is more salient. Investors can benefit most from focusing on the stockholdings of the top tercile performing active mutual funds.

Appendix: Tables and Figures

In this section, we confirm Berk and Binsbergen (2015) that fund gross value added (GVA) is a better measure of fund quality. In each style segment, Table II Panel A reports TNA in millions of dollars at quarter-end, and in parentheses, the average 4-factor gross alpha compounded over three months in the quarter. The $Size\_BTM$ column reports TNA, and in parentheses the 4-factor gross alpha, both averaged across momentum quintiles, on stocks sorted by size and book-to-market. Table II Panel B reports, for each style segment, the average quarterly gross value-added (GVA) in millions of dollars which is monthly GVA summed across three months in the quarter.

< Insert Appendix Table II here. >

Corroborating prior results on stock and fund ownerships, Online Appendix Figure I Panel A confirms that TNA increases across size quintiles. Figure I Panel B graphs 4-factor gross alpha and GVA against TNA. Gross alpha decreases with fund size. From the $Size\_BTM$ column in Table II Panel B, gross alpha declines from a high of 0.85% to a low of 0.57% as TNA grows from a low of \$482 million to a high of \$846 million. Further, GVA is concave in fund size. From the $Size\_BTM$ column Table II Panel B, GVA rises from an opening low of \$3.72 million on TNA of \$554 million to a peak high of \$4.85 million on TNA of \$770 million, then falling to a closing low of -\$0.53 million on TNA of \$846 million. GVA is also higher on low momentum stocks across all size and book-to-market quintiles. Lastly, GVA is higher on low book-to-market (growth) than high book-to-market (value) stocks across all size and momentum quintiles. The GVA gap of \$0.80 (=\$3.72-\$2.92) million is highest at the smallest size quintile and decreases to \$0.14 (=-\$0.39+\$0.53) million at the largest size quintile.

< Insert Online Appendix Figure I here. >

Using the quarterly distributions of gross alphas averaged across momentum quintiles on stocks sorted by size and book-to-market quintiles, we estimate a log-linear regression of gross alpha on the natural log of TNA controlling for quarter fixed effects. Results are reported in Column 1 in Appendix Table II Panel C. Our coefficient on $ln(TNA)$ of -0.0049, which is statistically significant at the 1% level, reflects a reduction in quarterly gross alpha of 49 bps on a 1% increase in TNA. In Table 12 of Zhu (2018), her estimated coefficient of -0.0020 indicates a 1% increase in TNA reduces monthly gross alpha by 20 bps.

We also estimate a quadratic log-linear regression with suppressed intercept of GVA on $ln(TNA)$ and its square, controlling for quarter fixed effects. Results are reported in Column 2 of Appendix Table II Panel C. The coefficients of 14.115 and -2.458, which are statistically significant at the 1% level, confirms a concave relationship between GVA and TNA. From the first-order condition, the predicted maximum GVA of \$7.1 million is attained at TNA of \$17.7 ($=exp(0.5 * 14.115/2.458)$ million.

< Insert Online Appendix Figure II here. >

Online Appendix Figure II graphs the frequency distribution of GVA and gross alpha against GVA. The predicted maximum GVA of \$7.1 million is realized at an average gross alpha of 1.51% and TNA of \$708 million. At average gross alpha of 0.05% and TNA of \$682 million, realized GVA is \$0.1 million. As shown in Appendix Table II Panel C, 31.2% of funds with negative gross alpha are overfunded, and 25.2% of funds with higher than predicted maximum GVA and significant positive gross alphas are underfunded. Compared to Table 7 in Zhu (2018), we have a lower percentage of overfunded funds and higher percentage of moderately funded funds. In contrast to Zhu (2018), our results are based on average fund holdings across style segments of stocks sorted by size, book-to-market, and momentum rather than average holdings across funds which do not take differences in investment strategies across funds into account.

In summary, in Appendix Table II Panel A, the average fund in the smallest size quintile has lower TNA under management but higher gross alpha. Average fund size rises, and gross alpha falls, with increasing size quintiles. In Appendix Table II Panel B and Online Appendix Figure I, breadth and depth of ownership affects GVA, and managerial skill exhibits diminishing returns to scale. In Table Appendix II Panel C and Online Appendix Figure II, GVA is negatively related to TNA for overfunded funds, and positively related, for underfunded funds. Moreover, gross alpha is negative for overfunded funds, and highly positive, for underfunded funds. TNA and gross alpha do not correlate with GVA when funds are over- or under-funded.

References

Ali, A., C. Durtschi, B. Lev, and M. Trombley. 2004. Changes in Institutional Ownership and Subsequent Earnings Announcement Abnormal Returns. *Journal of Accounting, Auditing & Finance* 19, 221-248.

Amihud, Y. and R. Goyenko. 2013. Mutual Fund's R2 as Predictor of Performance. *Review of Financial Studies* 26, 667-694.

Baker, M., L. Litov, J. Wachter, and J. Wurgler. 2010. Can Mutual Fund Managers Pick Stocks? Evidence from their Trades Prior to Earnings Announcements. *Journal of Financial and Quantitative Analysis* 45, 1111-1131.

Berk, J., and J. van Binsbergen. 2015. Measuring Skill in the Mutual Fund Industry. *Journal of Financial Economics* 118, 1-20.

Berk, J. and R. Green, 2004. Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy* 112: 1269–1295.

Bernhardt, Dan, and Eric Hughson. 2015. Splitting Orders. *The Review of Financial Studies* 10, 69-101.

Brennan, M., T. Chordia, and A. Subrahmanyam. 1998. Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns. *Journal of Financial Economics* 49, 345-373.

Brown, D. and S. Davies. 2017. Moral Hazard in Active Asset Management. *Journal of Financial Economics* 125, 311-325.

Brown, Nerissa C.; Kelsey D. Wei and Russ Wermers. 2014. Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices. *Management Science* 60, 1-20.

Cao, B., R. Wermers. 2018. Mutual Fund Links (MFLINKS) Update of Nov 2018. https://wrds-www.wharton.upenn.edu/documents/1018/MFLINKS_2018.pdf

Carhart, M. 1997. On Persistence in Mutual Fund Performance. Journal of Finance 52, 57-82.

Chen, J., H. Hong, M. Huang, and J. D. Kubik, 2004. Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization. *American Economic Review* 94: 1276–1302

Chen, J., H. Hong, and J.C Stein. 2002. Breadth of Ownership and Stock Returns. *Journal of Financial Economics* 66, 171-205.

Chevalier, J. and G. Ellison. 1999. Are some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance. *Journal of Finance* 54, 875-899.

Cohen, R., J. Coval, and L. Pastor, 2005. Judging Fund Managers by the Company that They Keep. *Journal of Finance* 60:1057-1096.

Cremers, M., J. Fulkerson, and T.B. Riley. 2019. Challenging the Conventional Wisdom on Active Management: A Review of the Past 20 Years of Academic Literature on Actively Managed Mutual Funds. *Financial Analysts Journal* 75, 8-35.

Cremers, M. and A. Pareek, 2016. Patient Capital Outperformance: The Investment Skill of High Active Share Managers Who Trade Infrequently. *Journal of Financial Economics*, 122:288-306.

Cremers, M. and A. Petajisto, 2009. How Active is Your Fund Manager? A New Measure that Predicts Performance. *Review of Financial Studies* 22:3329-3365.

Dalio, R. 2019. Work Principle 5; Believability Weight Your Decision Making. LinkedIn .https://www.linkedin.com/pulse/work-principle-5-believability-weight-your-decision-making-ray-dalio.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997. Measuring Mutual Fund Performance with Characteristic-based Benchmarks. *Journal of Finance* 52:1035-1058.

Daniel, K. and S. Titman. 2006. Market Reactions to Tangible and Intangible Information. *Journal of Finance* 61, 1605-1643.

Del Guercio, D. and P. Tkac, 2002. The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds vs. Pension Funds. *Journal of Financial and Quantitative Analysis* 37: 523–557.

Diether, K., C. Malloy, and A. Scherbina, 2002. Differences of Opinion and the Cross-section of Stock Returns. *Journal of Finance* 57:2113-2141.

Doshi, H., R. Elkamhi, and M. Simutin. 2015. Managerial Activeness and Mutual Fund Performance. *Review of Asset Pricing Studies* 5, 156-184.

Elton, E., M. Gruber, and C. Blake. 2003. Incentive Fees and Mutual Funds. *Journal of Finance* 58, 779-804.

Elton, E., M. Gruber, and C. Blake, 2012. Does Mutual Fund Size Matter? The Relationship between Aize and Performance. *Review of Asset Pricing Studies* 2: 31–55.

Evans, R., M. Prado, and R. Zambrana. 2020. Competition and Cooperation in Mutual Fund Families. *Journal of Financial Economics* 136, 168-188.

Fama, E. and K. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E. and K. French. 1995. Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance* 50, 131-155.

Fama, E. and J. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81, 607-636.

Frazzini, A., J. Friedman, and L. Pomorski. 2016. Deactivating Active Share. *Financial Analysts Journal* 72, 14-21.

Frazzini, Andrea and Owen A. Lamont. 2008. Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns. *Journal of Financial Economics* 88, 299-322.

French, K. 2008. Presidential Address: The Cost of Active Investing. *Journal of Finance* 63, 1537-1573.

Gârleanu, N. and L. Pedersen. 2018. Efficiently Inefficient Markets for Assets and Asset Management. *Journal of Finance* 73, 1663-1712.

Gaspar, J.; M. Massa, and P. Matos. 2005. Shareholder Investment Horizons and the Market for Corporate Control. *Journal of Financial Economics* 76, 135-165.

Gompers. P. and A. Metrick, 2001. Institutional Investors and Equity Prices. *Quarterly Journal of Economics* 116:229–259.

Grossman, S. and J. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70, 393-408.

Jegadeesh, N. 1990. Evidence of Predictable Behavior of Security Returns. *Journal of Finance* 45, 881-898.

Jiang, H. and Z. Sun, 2014. Dispersion in Beliefs among Active Mutual Funds and the Cross-section of Stock Returns. *Journal of Financial Economics* 114:341-365.

Jiang, H., M. Verbeek, and Y. Wang, 2020. Information Content when Mutual Funds Deviate from Benchmarks. *Management Science* 60:2038-2053

Jiang, H. and L. Zheng, 2018. Active Fundamental Performance. *Review of Financial Studies* 31:4688-4719.

Hou, K., C. Xue, and L. Zhang. 2020. Replicating Anomalies. *Review of Financial Studies* 33, 2019-2133.

Hunter, D.; E. Kandel; S. Kandel and R. Wermers. 2014. Mutual Fund Performance Evaluation with Active Peer Benchmarks. *Journal of Financial Economics* 112, 1-29.

Jones, S.; D. Lee and E. Weis. 1999. Herding and Feedback Trading by Different Types of Institutions and the Effects on Stocks. Indiana University-Indianapolis, Kennesaw State University, and Merrill Lynch. Working Paper.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp, 2014. Time-varying Fund Manager Skill. *Journal of Finance* 69:1455-1484.

Kacperczyk, M. and A. Seru, 2007. Fund Manager Use of Public Information: New Evidence on Managerial Skills. *Journal of Finance* 62:485–528.

Kacperczyk, M., C. Sialm, and L. Zheng, 2005. On the Industry Concentration of Actively Managed Equity Mutual Funds. *Journal of Finance* 60:1983–2011.

Kacperczyk, M., C. Sialm and L. Zheng. 2008. Unobserved Actions of Mutual Funds. *Review of Financial Studies* 21, 2379-2416.

Kyle, A., 1985. Continuous Auction and Insider Trading. *Econometrica* 53, 1315-1335.

Lakonishok, Josef; Andrei Shleifer and Robert W. Vishny. 1992. The Impact of Institutional Trading on Stock Prices. *Journal of Financial Economics* 32, 23-43.

Lan, C, and F. Moneta, and Wermers, R. 2019. Holding Horizon: A New Measure of Active Investment Management. SSRN: https://ssrn.com/abstract=2517215.

Mclean, R. and P. Jeffrey 2016. Does Academic Research Destroy Stock Return Predictability? *Journal of Finance* 71, 5-32.

Mamaysky, H., M. Spiegel, and H. Zhang. 2007. Improved Forecasting of Mutual Fund Alphas and Betas*. *Review of Finance* 11, 359-400.

Miller, E. 1977. Risk, Uncertainty, and Divergence of Opinion. *Journal of Finance* 32, 1151-1168.

Newey, W. and K. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

O'Hara, Maureen, Chen Yao and Mao Ye. 2014. What's Not There: Odd Lots and Market Data. *The Journal of Finance* 69, 2199-2236.

Pástor, L., R. Stambaugh, and L. Taylor. 2015. Scale and Skill in Active Management. *Journal of Financial Economics* 116, 23-45.

Pastor, L.; R. Stambaugh and L. Taylor. Fund Tradeoffs. *Journal of Financial Economics*. forthcoming.

Petajisto, A. 2013. Active Share and Mutual Fund Performance. *Financial Analysts Journal* 69:73–93.

Sensoy, B. 2008. Performance Evaluation and Self-designated Benchmark Indexes in the Mutual Fund Industry. *Journal of Financial Economics* 92: 25-39.

Shawky, Hany A., and David M. Smith, 2005, Optimal Number of Stock Holdings in Mutual Fund Portfolios Based on Market Performance, Financial Review 40, 481-495.

Sias, R., 2004. Institutional Herding. *Review of Financial Studies* 17:165–206.

Wermers, R. 2000. Mutual Fund Performance: An Empirical Decomposition into Stock-picking Talent, Style, Transaction Costs, and Expenses. *Journal of Finance* 55:1655-1695.

Wermers, R. 2003. Is Money Really 'Smart'? New Evidence on the Relation between Mutual Fund Flows, Manager Behavior, and Performance Persistence.
SSRN: https://ssrn.com/abstract=414420or http://dx.doi.org/10.2139/ssrn.414420

Wermers, Russ; Tong Yao and Jane Zhao. 2012. Forecasting Stock Returns through an Efficient Aggregation of Mutual Fund Holdings. *The Review of Financial Studies* 25, 3490-3529.

Yan, X. 2008. Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance. *Journal of Financial and Quantitative Analysis* 43: 741-768.

Zhu, M. 2018. Informative Fund Size, Managerial Skill, and Investor Rationality. *Journal of Financial Economics* 130: 114-134.

Table I: Distributions of Stock and Fund Characteristics by Style Segments

Table reports median style segment characteristics. Cutoffs from annual Daniel, Grinblatt, Titman and Wermers (1997) sorts of stocks into quintiles by size, industry-adjusted book-to-market, and momentum are used to assign stocks into 125 style segments each quarter over our 72-quarter sample period 2000-2017. For each style segment we determine the number of stocks owned. In addition, for each fund in a style segment, we compute the number of stocks owned as a percentage of CRSP stocks in the style. Median number and percentages are reported in Panel A. Panel B reports the number of funds who own the stock., and for each stock in the style segment, the number of funds who own the stock as a percentage of the number of funds who own stocks in the style segment. Median number and percentages are reported in parentheses.

		PANEL A: Number of Stocks Held by Funds (Stock Ownership)						**Panel B: Number of Funds (Fund Ownership)**					
SIZE	**BTM**	**MOMENTUM**						**MOMENTUM**					
		1	2	3	4	5	***Size_BTM***	1	2	3	4	5	***Size_BTM***
1	1	30 (4.6)	48 (3.6)	54 (3.6)	54 (3.5)	55 (3.6)	48 (3.8)	149 (9.3)	214 (7.8)	246 (8.1)	272 (8.1)	232 (8.1)	223 (8.3)
1	2	39 (3.7)	53 (3.4)	57 (3.3)	58 (3.2)	59 (3.4)	53 (3.4)	195 (8.5)	244 (8.2)	243 (8.6)	243 (8.1)	235 (7.9)	232 (8.3)
1	3	40 (3.9)	60 (3.3)	63 (3.2)	60 (3.1)	59 (3.3)	56 (3.4)	168 (9.3)	222 (9.0)	230 (8.8)	250 (9.0)	247 (8.8)	223 (9.0)
1	4	42 (4.0)	53 (3.5)	61 (3.4)	52 (3.3)	60 (3.7)	53 (3.5)	163 (9.4)	197 (9.0)	198 (9.4)	194 (9.0)	212 (8.7)	193 (9.1)
1	5	34 (4.8)	45 (3.5)	50 (3.3)	52 (3.3)	52 (3.6)	46 (3.7)	132 (11.0)	169 (8.5)	167 (8.0)	192 (7.9)	200 (8.2)	172 (8.7)
2	1	23 (6.5)	24 (7.0)	23 (8.3)	24 (8.2)	24 (8.2)	23 (7.5)	310 (12.0)	312 (11.9)	333 (12.8)	320 (12.8)	320 (13.4)	319 (12.5)
2	2	23 (7.3)	24 (7.6)	23 (7.9)	24 (8.0)	23 (8.5)	23 (7.7)	303 (12.2)	328 (12.5)	325 (12.8)	330 (12.8)	316 (13.7)	320 (12.8)
2	3	24 (7.1)	24 (7.6)	25 (7.6)	24 (7.9)	24 (7.9)	24 (7.5)	306 (13.0)	331 (13.1)	323 (13.4)	327 (14.2)	330 (13.6)	323 (13.5)
2	4	23 (6.9)	25 (7.9)	25 (7.8)	24 (7.7)	23 (8.6)	24 (7.6)	297 (13.4)	308 (14.9)	323 (14.7)	321 (14.5)	320 (15.5)	314 (14.6)
2	5	23 (6.1)	24 (6.7)	24 (7.3)	24 (7.2)	23 (7.7)	24 (6.8)	277 (13.9)	307 (13.5)	316 (14.4)	328 (14.3)	314 (15.1)	308 (14.2)
3	1	16 (7.4)	17 (9.0)	17 (9.8)	17 (10.4)	16 (11.7)	17 (9.2)	410 (13.1)	425 (13.4)	429 (14.1)	433 (15.1)	386 (16.7)	416 (14.5)
3	2	16 (7.2)	17 (9.9)	17 (9.8)	17 (10.6)	16 (11.3)	17 (9.4)	404 (13.2)	419 (14.1)	424 (14.3)	428 (15.0)	407 (16.0)	416 (14.5)
3	3	17 (7.2)	17 (8.7)	16 (9.8)	17 (9.3)	16 (11.2)	17 (8.7)	406 (12.9)	440 (13.5)	418 (14.7)	430 (14.6)	409 (15.8)	420 (14.3)
3	4	16 (7.6)	17 (7.8)	17 (9.3)	16 (9.8)	16 (11.0)	16 (8.6)	360 (14.1)	419 (13.7)	408 (14.9)	429 (15.4)	405 (16.0)	404 (14.8)
3	5	16 (6.8)	17 (8.4)	16 (8.2)	17 (8.9)	16 (9.1)	16 (8.1)	359 (14.5)	396 (14.9)	409 (15.0)	420 (15.1)	418 (15.3)	400 (15.0)
4	1	14 (7.5)	15 (7.1)	14 (7.5)	14 (10.0)	14 (11.4)	14 (8.0)	455 (14.5)	499 (14.1)	482 (14.8)	536 (15.5)	525 (16.5)	499 (15.1)
4	2	14 (8.0)	14 (8.0)	15 (7.7)	14 (9.8)	14 (11.0)	14 (8.4)	456 (14.8)	501 (14.6)	527 (13.9)	531 (15.4)	525 (16.9)	508 (15.1)
4	3	14 (7.2)	14 (7.1)	15 (7.8)	14 (8.5)	13 (11.0)	14 (7.7)	445 (14.6)	488 (14.6)	513 (14.7)	522 (15.3)	532 (16.5)	500 (15.1)
4	4	14 (7.7)	14 (7.9)	14 (9.1)	14 (9.1)	14 (10.5)	14 (8.4)	434 (15.7)	476 (14.6)	515 (15.5)	493 (15.2)	528 (16.1)	489 (15.4)
4	5	14 (7.7)	14 (8.1)	14 (8.7)	14 (8.4)	14 (10.8)	14 (8.3)	387 (16.9)	430 (15.8)	433 (16.1)	445 (16.2)	494 (17.4)	438 (16.5)
5	1	13 (17.2)	13 (16.4)	13 (19.0)	13 (18.5)	13 (17.6)	13 (17.8)	622 (19.2)	696 (19.5)	689 (22.4)	667 (21.6)	648 (22.2)	664 (21.0)
5	2	12 (16.9)	13 (18.0)	13 (17.8)	13 (17.9)	13 (17.3)	13 (17.6)	673 (19.6)	687 (20.8)	703 (20.3)	721 (21.1)	677 (21.2)	692 (20.6)
5	3	13 (16.5)	13 (16.0)	13 (18.1)	13 (16.5)	13 (17.0)	13 (16.8)	647 (19.1)	687 (19.3)	685 (19.9)	703 (20.0)	672 (20.9)	679 (19.8)
5	4	12 (16.5)	13 (16.7)	13 (16.3)	13 (16.8)	12 (16.4)	13 (16.6)	602 (19.9)	617 (19.9)	670 (19.9)	707 (20.2)	635 (21.5)	646 (20.3)
5	5	12 (16.0)	13 (17.2)	13 (16.2)	13 (16.8)	12 (16.4)	13 (16.6)	526 (20.2)	582 (20.5)	578 (20.7)	595 (20.7)	588 (21.7)	574 (20.7)

Table II: Summary Statistics

Table reports stock summary statistics on variables used in the paper. The number of stock-months is 191,274 in the sample. Excess market return is the monthly return in excess of the value weighted CRSP return compounded over a quarter. DGTW return is the monthly return minus the average return on stocks in DGTW segment style k to which stock i belongs compounded over a quarter. 4-Factor adjusted alpha is the daily alpha estimated from time-series regressions of daily stock returns on Fama and French (1992) market, SMB, HML, and Carhart (1997) UMD factors each month compounded over a quarter. SUE is earnings per share minus median analyst forecast made earlier than the earnings announcement date but no more than 90 days in advance scaled by stock price at the end of quarter. Breadth is $ln(N+1)$, where N is the number of actively managed mutual funds with non-zero holdings of stock i in style segment k. Active MF ownership is the percentage of total shares outstanding of stock i owned by actively managed mutual funds j at the end of quarter q. Quarterly change in breadth and active MF ownership are computed as change in breadth and active MF ownership from the prior quarter. Dispersion index of active holding is the standard deviation of active holdings across all funds j with non-zero holdings in style segment k divided by the absolute value of the mean active holding. Market capitalization is the product of closing price and total shares outstanding of stock i at the end of the quarter q expressed in millions of dollars. Book-to-market is book equity to shareholders' equity following Daniel and Titman (2006). Prior year return in month t at quarter end is the cumulative monthly return over the prior 12 months starting from $t-2$ and ending in $t-13$. CRSP turnover is the total trading volume reported by CRSP summed across all 3 months in the quarter as a percentage of total shares outstanding where trading volume is adjusted following French (2008). Idiosyncratic volatility is the standard deviation of residuals from time series regressions of daily stock returns on Fama French (1992) market, SMB and HML factors each quarter. Market beta is the sum of the coefficients on contemporaneous and five lags of market excess returns estimated from time series regressions of daily stock excess returns on daily contemporaneous and five lags of market excess returns each quarter following Jiang and Sun (2014). As in Berk and Binsbergen (2015), we proxy fund performance by gross-value added (GVA). Monthly gross value-added is the product of current month 4-factor alpha and prior month end TNA. TNA is deflated by the monthly seasonally adjusted CPI index (1982-1984=100) constructed by the U.S. Bureau of Labor Statistics (BLS). Quarterly gross value added is monthly gross value added summed across months in the quarter. Fund holding of stock i by fund j is the market value of stock i owned by fund j as a percentage of all stock holdings of fund j in style segment k. Portfolio formation is described in Table I. Peer group holding is the market value of stock i owned by all actively managed mutual funds j as a percentage of the market value of all stocks in style segment k owned by all actively managed mutual funds j. Active holding is the deviation of fund from peer group holding. Active holding turnover is the difference in active holding between current and four quarters prior. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. In $\theta_i^s\_Mgmt\ Fee$, management fees proxy for GVA. The cross-product of active holding turnover and GVA summed across funds is used to identify active turnover – the patience and conviction of fund managers. Conviction quality $Turn \equiv \theta_i^\pi$ is the odds ratio of a stock's relative percentile rank, θ_i^π, on active turnover.

	Mean	Std Dev	Percentile 10th	20th	50th	80th	90th
Excess Market Return (%)	0.70	10.81	-11.11	-6.30	0.32	7.37	12.64
DGTW Return (%)	-0.06	10.45	-11.65	-6.87	-0.29	6.44	11.48
4-Factor Alpha (%)	0.59	12.45	-12.43	-7.20	-0.06	7.50	13.89
Market Capitalization ($million)	5,118	21,384	69	149	672	3,390	9,050
Book-to-Market	0.745	0.797	0.187	0.287	0.575	1.018	1.380
Prior Year Return	0.138	0.442	-0.377	-0.155	0.137	0.423	0.634
CRSP Turnover	0.435	0.411	0.065	0.122	0.319	0.654	0.940
Idiosyncratic Volatility	0.023	0.015	0.009	0.012	0.019	0.031	0.042
Market Beta	1.177	1.172	-0.070	0.325	1.045	1.950	2.618
$\Delta Breadth$	0.008	0.173	-0.150	-0.086	0.000	0.102	0.167
$\Delta Active\ MF\ Ownership$	0.000	0.029	-0.016	-0.007	0.000	0.008	0.016
$Dispersion\ Index\ in\ Active\ Holdings$	0.137	0.084	0.011	0.057	0.139	0.210	0.245
$Breadth$	3.606	0.918	2.303	2.708	3.689	4.407	4.745
$Active\ MF\ Ownership$	0.108	0.120	0.017	0.032	0.080	0.158	0.208
$Investment\ Quality,\ \theta_i^s$	4.572	12.760	0.111	0.250	1.000	3.997	8.986
$IQ,\ \theta_i^s\_Mgmt\ Fee$	4.572	12.760	0.111	0.250	1.000	3.997	8.986
$IQ,\ \theta_i^s\_Industry\ Concentration$	4.572	12.760	0.111	0.250	1.000	3.997	8.986
$Conviction\ Quality\ \theta_i^\pi$	4.572	12.760	0.111	0.250	1.000	3.997	8.986
$Investment\ Quality,\ \theta_i^{s^2}$	183.72	1,076.51	0.012	0.063	1.000	15.974	80.742
$IQ,\ \theta_i^{s^2}\_Mgmt\ Fee$	183.72	1,076.51	0.012	0.063	1.000	15.974	80.742
$IQ,\ \theta_i^{s^2}\_Industry\ Concentration$	183.72	1,076.51	0.012	0.063	1.000	15.974	80.742

Table III: Persistence in Stock Investment Quality

This table reports on the persistence of stock investment quality through lead four quarters. Stock investment quality is defined in Table II and in the Appendix. In Panel A, we sort stocks into deciles on investment quality at the end of each quarter. In each decile, we report the average stock investment quality in lead two, four, eight, and twelve quarters, as well as the spread in investment quality between the top and bottom deciles and associated t-statistics. In Panel B, we report a transition matrix. We sort stocks by investment quality into quintiles at the end of each quarter and compute the percentage of stocks that remain or change to another quintile in the subsequent quarter. Panel C presents summary statistics on investment quality in the selected percentiles between the 5th and the 95th percentiles. Superscripts a,b,c denote two-tailed tests of statistical significance at the 10%, 5%, and 1% levels.

Panel A	Investment Quality			
	Lead 2 Qtr	Lead 4 Qtr	Lead 8 Qtr	Lead 12 Qtr
Low	1.50	1.40	1.24	1.16
2	1.06	1.01	1.05	1.18
3	1.04	1.19	1.34	1.43
4	1.35	1.49	1.65	1.73
5	1.62	1.79	2.02	2.15
6	2.03	2.30	2.53	2.64
7	2.69	2.91	3.28	3.44
8	3.83	4.16	4.43	4.66
9	6.51	6.74	6.98	7.03
High	23.30	21.37	19.03	17.30
High-Low	21.80[c]	19.97[c]	17.80[c]	16.14[c]
t-stat	97.51	86.56	74.01	65.82

Panel B		Investment Quality Transition Matrix				
		Current Quarter End				
		1	2	3	4	5
Prior Quarter End	1	75.74%	13.88%	4.60%	2.74%	3.05%
	2	12.59%	55.39%	21.22%	7.13%	3.66%
	3	4.45%	20.24%	49.20%	20.79%	5.32%
	4	2.75%	7.15%	19.90%	54.03%	16.17%
	5	3.31%	3.81%	5.38%	15.63%	71.86%

Panel C	Investment Quality Percentile Rank p	Active Holdings (%)	GVA	$\rho\left(\begin{matrix}AH,\\GVA\end{matrix}\right)$	MCAP	Book-to-Market
	0.05	-0.002	4.55	-0.002	890	1.125
	0.10	-0.011	5.00	-0.017	2,754	0.992
	0.20	-0.033	5.92	-0.025	4,621	0.740
	0.50	0.007	5.59	-0.028	4,042	0.703
	0.80	0.015	5.90	0.011	6,442	0.649
	0.90	0.013	5.77	0.049	8,848	0.646
	0.95	0.005	5.76	0.092	9,493	0.652

Table IV: Alternative Empirical Stock Return Predictors: Herding, Delta Ownership, Delta Breadth and Flow Effect

This table compares stock investment quality with four other empirical measures used to forecast future stock returns. Table reports Fama-Macbeth (1973) regression results of lead quarter buy-and-hold stock returns on stock investment quality, herding (Brown, Wei, and Wermers, 2014), flow (Frazzini and Lamont, 2007), delta fund ownership and delta breadth (Chen, Hong, and Stein, 2002). In Fama-Macbeth regressions, serial correlation in error terms are corrected using a Newey-West estimator with one-quarter lag. p-values are reported in parentheses. Superscript a, b, c denote statistical significance at 10%, 5%, and 1% levels. Unreported estimated coefficients on control variables are reported in an Appendix Table IV.

	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ \equiv \theta_i^s$	1.171^{c}	1.037^{c}	0.920^{c}	0.819^{c}	1.187^{c}	1.042^{c}	0.935^{c}	0.826^{c}	1.040^{c}	0.903^{c}	0.807^{c}	0.711^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Herding	-0.145^{b}	-0.050	0.002	0.011	-0.183^{c}	-0.085^{a}	-0.022	-0.013	-0.113	-0.045	-0.010	0.020
	(0.039)	(0.288)	(0.951)	(0.755)	(0.006)	(0.068)	(0.528)	(0.705)	(0.142)	(0.498)	(0.854)	(0.697)
Flow	0.150	0.159	0.192^{b}	0.201^{b}	0.142	0.152^{a}	0.195^{b}	0.205^{c}	0.146	0.154^{a}	0.143^{b}	0.151^{b}
	(0.200)	(0.101)	(0.036)	(0.023)	(0.170)	(0.084)	(0.014)	(0.010)	(0.144)	(0.058)	(0.043)	(0.034)
$\Delta Breadth$	0.383^{b}	0.217^{a}	0.060	-0.008	0.351^{b}	0.190	0.047	-0.015	0.300^{b}	0.178^{b}	0.014	-0.045
	(0.015)	(0.065)	(0.527)	(0.920)	(0.020)	(0.117)	(0.618)	(0.841)	(0.017)	(0.047)	(0.841)	(0.434)
$\Delta Active\ MF\ Ownership$	-0.145^{a}	-0.131^{b}	-0.054	-0.034	-0.155^{b}	-0.135^{b}	-0.083^{a}	-0.068	-0.173^{b}	-0.157^{c}	-0.051	-0.029
	(0.050)	(0.021)	(0.254)	(0.452)	(0.035)	(0.025)	(0.056)	(0.123)	(0.040)	(0.007)	(0.250)	(0.541)
$IQ^2 \equiv \theta_i^{s\,2}$	-0.867^{c}	-0.802^{c}	-0.710^{c}	-0.658^{c}	-0.898^{c}	-0.826^{c}	-0.752^{c}	-0.690^{c}	-0.828^{c}	-0.725^{c}	-0.655^{c}	-0.585^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Holding Dispersion Index	-0.010	-0.067^{a}	-0.049	-0.018	-0.041	-0.064^{a}	-0.046	-0.017	0.034	-0.050	-0.035	0.008
	(0.822)	(0.075)	(0.182)	(0.547)	(0.386)	(0.089)	(0.163)	(0.525)	(0.494)	(0.197)	(0.309)	(0.792)
NOBS	178,535	175,711	172,920	170,167	178,424	175,495	172,583	169,718	180,929	178250	175,505	172,798
R^2	0.066	0.071	0.076	0.078	0.032	0.037	0.040	0.043	0.031	0.040	0.046	0.053

Table V: Excess Returns on One-Way Sort by Stock Investment Quality

Table reports the average value-weighted buy-and-hold returns in the month and quarters following the formation of quintile portfolios on stocks sorted by investment quality. Monthly and average quarterly returns are expressed in percent. In each quarter, we estimate the percentage of total dollar holdings in a style segment allocated by a fund and its peers to the same stock. Active holding is the difference between fund and peer group holding of the same stock in a style segment. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. Other variable definitions can be found in Table II. Superscripts [a,b,c] denote two-tailed tests of statistical significance at the 10%, 5%, and 1% levels.

	Lead 1 Mo	**Lead 1 Qtr**	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**
			EXCESS MARKET RETURN		
Low	-0.406	-0.774	-1.289	-1.392	-1.329
2	-0.199	-0.336	-0.753	-0.928	-0.982
3	0.026	-0.026	-0.449	-0.517	-0.564
4	0.123	0.548	0.033	-0.139	-0.277
High	0.269	0.759	0.403	0.277	0.102
High-Low	**0.675**[c]	**1.533**[c]	**1.692**[c]	**1.668**[c]	**1.431**[c]
***t*-stat**	2.75	3.05	5.74	6.10	5.73
			DGTW RETURN		
Low	-0.487	-1.013	-1.407	-1.498	-1.463
2	-0.274	-0.432	-0.696	-0.845	-0.877
3	-0.081	-0.199	-0.587	-0.602	-0.614
4	0.017	0.314	-0.138	-0.317	-0.430
High	0.137	0.710	0.361	0.196	0.053
High-Low	**0.624**[c]	**1.723**[c]	**1.768**[c]	**1.694**[c]	**1.516**[c]
***t*-stat**	2.83	4.38	6.82	6.97	6.69
			4-FACTOR ALPHA		
Low	-0.097	-0.561	-0.770	-0.781	-0.846
2	0.087	0.214	-0.128	-0.299	-0.360
3	0.379	0.309	-0.100	-0.158	-0.143
4	0.335	0.717	0.167	-0.053	-0.184
High	0.575	0.975	0.611	0.392	0.232
High-Low	**0.672**[b]	**1.537**[c]	**1.381**[c]	**1.173**[c]	**1.078**[c]
***t*-stat**	2.53	3.60	4.22	4.11	4.16

Table VI: Excess Returns on Double-Sorts by Market Capitalization and Stock Investment Quality

Table reports the average value-weighted buy-and-hold returns in the quarters following portfolio formation of stocks double-sorted first into terciles by market capitalization using NYSE stocks to establish breakpoints, and second, by investment quality. Average quarterly returns are expressed in percent. Portfolio formation is described in Table I. In each quarter, we estimate the percentage of total dollar holdings in a style segment allocated by a fund and its peers to the same stock. Active holding is the difference between fund and peer group holding of the same stock in a style segment. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. Other variable definitions can be found in Table II. Superscripts [a,b,c] denote two-tailed tests of statistical significance at the 10%, 5%, and 1% levels.

		EXCESS MARKET RETURN			DGTW RETURN			4-FACTOR ALPHA		
		Market Capitalization			Market Capitalization			Market Capitalization		
		Small	Mid	Large	Small	Mid	Large	Small	Mid	Big
Lead 1 Quarter										
Stock Selection Quality	Low	0.225	-0.093	-0.637	-1.031	-1.057	-0.603	0.502	-0.305	-0.093
	2	0.079	0.595	-0.241	-1.192	-0.425	-0.310	-0.137	0.217	0.366
	3	1.323	1.089	0.191	0.171	0.150	0.210	1.001	0.567	0.601
	4	1.822	1.090	0.316	0.731	0.156	0.319	1.360	0.584	0.823
	High	2.050	1.503	1.009	0.876	0.515	1.004	1.171	0.788	1.148
	Hi-Lo	1.825[a]	1.596[c]	1.646[c]	1.907[c]	1.572[c]	1.607[c]	0.670[b]	1.094[c]	1.241[c]
	***t*-stat**	1.931	2.777	5.362	6.711	5.827	4.656	2.142	2.972	3.023
Lead 2 Quarter										
Stock Selection Quality	Low	-0.850	-0.722	-0.968	-1.928	-1.582	-0.769	-0.503	-0.880	-0.253
	2	-0.937	-0.116	-0.660	-1.995	-1.036	-0.615	-1.052	-0.459	-0.103
	3	0.027	0.273	-0.089	-0.990	-0.566	-0.137	-0.139	-0.034	0.328
	4	0.502	0.357	0.101	-0.463	-0.455	0.127	0.165	-0.060	0.512
	High	0.487	0.508	0.569	-0.514	-0.365	0.576	-0.120	-0.042	0.691
	Hi-Lo	1.337[c]	1.230[c]	1.538[c]	1.414[c]	1.217[c]	1.345[c]	0.383	0.839[c]	0.944[c]
	***t*-stat**	2.931	2.991	7.966	6.136	6.093	9.293	1.587	4.218	4.029
Lead 3 Quarter										
Stock Selection Quality	Low	-1.099	-0.892	-1.038	-2.142	-1.716	-0.810	-0.779	-1.094	-0.263
	2	-1.310	-0.360	-0.767	-2.303	-1.240	-0.685	-1.457	-0.655	-0.206
	3	-0.365	0.086	-0.207	-1.348	-0.755	-0.233	-0.606	-0.233	0.167
	4	0.035	0.093	-0.037	-0.875	-0.695	-0.021	-0.324	-0.250	0.309
	High	0.002	0.173	0.415	-0.972	-0.651	0.399	-0.594	-0.365	0.439
	Hi-Lo	1.102[c]	1.064[c]	1.453[c]	1.170[c]	1.065[c]	1.209[c]	0.185	0.729[c]	0.702[c]
	***t*-stat**	3.009	4.551	9.080	5.726	6.362	10.314	0.861	4.328	3.606
Lead 4 Quarter										
Stock Selection Quality	Low	-1.258	-1.010	-1.062	-2.206	-1.745	-0.822	-0.936	-1.172	-0.343
	2	-1.574	-0.446	-0.779	-2.428	-1.231	-0.678	-1.702	-0.713	-0.154
	3	-0.572	-0.029	-0.356	-1.463	-0.803	-0.385	-0.785	-0.331	0.026
	4	-0.180	-0.152	-0.162	-0.993	-0.865	-0.088	-0.510	-0.431	0.143
	High	-0.360	-0.076	0.325	-1.206	-0.801	0.301	-0.832	-0.578	0.363

Table VI (Cont.)	EXCESS MARKET RETURN			DGTW RETURN			4-FACTOR ALPHA		
	Market Capitalization			Market Capitalization			Market Capitalization		
Lead 4 Quarter	Small	Mid	Large	Small	Mid	Large	Small	Mid	Big
Hi-Lo	0.898[c]	0.934[c]	1.386[c]	1.000[c]	0.944[c]	1.123[c]	0.104	0.594[c]	0.706[c]
t-stat	2.819	4.813	9.510	5.381	6.406	10.739	0.526	4.092	4.041

Table VII: Forecast Returns: Investment Quality of Stocks

Table VI Panel A reports style and quarter fixed effects regressions of lead quarter buy-and-hold stock returns on selection quality and control variables, and in Panel C, Fama-Macbeth (1973) regressions. Panel B reports forecast returns at selected percentiles of selection quality between the 5th and the 95th percentiles.‡ indicates estimated coefficients from Panel A. Variable definitions can be found in Table II. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. In two-way fixed effects regressions, errors are clustered by style and quarter. In Fama-Macbeth regressions, serial correlation in error terms are corrected using a Newey-West estimator with one-quarter lag. p-values are reported in parentheses. Superscript a, b, c denote statistical significance at 10%, 5%, and 1% levels.

Panel A	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ \equiv \theta_i^s$	1.360c	1.232c	1.140c	1.018c	1.333c	1.190c	1.083c	0.948c	0.895c	0.851c	0.844c	0.766c
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ^2 \equiv \theta_i^{s2}$	-1.088c	-1.031c	-0.972c	-0.876c	-1.063c	-0.984c	-0.919c	-0.820c	-0.700c	-0.706c	-0.729c	-0.651c
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\Delta Breadth$	0.533c	0.248b	0.100	0.028	0.497c	0.241c	0.111	0.046	0.312c	0.231c	0.063	0.015
	(0.004)	(0.029)	(0.253)	(0.706)	(0.001)	(0.010)	(0.114)	(0.451)	(0.009)	(0.006)	(0.335)	(0.802)
$\Delta Active\ MF\ Ownership$	-0.062	-0.112a	-0.071	-0.044	-0.075	-0.099a	-0.071	-0.054	-0.044	-0.099a	-0.035	-0.016
	(0.372)	(0.060)	(0.231)	(0.366)	(0.264)	(0.055)	(0.174)	(0.222)	(0.512)	(0.053)	(0.496)	(0.711)
Holding Dispersion Index	-0.006	-0.058	-0.056	-0.028	-0.051	-0.072a	-0.062a	-0.033	-0.004	-0.072b	-0.052a	-0.011
	(0.906)	(0.116)	(0.102)	(0.320)	(0.296)	(0.057)	(0.064)	(0.231)	(0.933)	(0.034)	(0.091)	(0.663)
Ln(MCAP)	-0.832	-0.400	-0.294	-0.212	-1.006a	-0.464	-0.223	-0.063	-0.671	-0.440	-0.444	-0.399
	(0.129)	(0.327)	(0.366)	(0.438)	(0.051)	(0.191)	(0.418)	(0.770)	(0.322)	(0.431)	(0.323)	(0.323)
Ln(Book-to-Market)	0.674	0.824b	0.792c	0.791c	0.111	0.283	0.330a	0.370b	-0.024	0.231	0.357b	0.432c
	(0.176)	(0.015)	(0.002)	(0.001)	(0.739)	(0.201)	(0.058)	(0.014)	(0.915)	(0.177)	(0.021)	(0.002)
Prior Year Return	-0.590	-0.541	-0.511a	-0.500a	-0.313	-0.247	-0.241	-0.235	0.116	-0.012	-0.066	-0.114
	(0.219)	(0.166)	(0.098)	(0.059)	(0.309)	(0.233)	(0.153)	(0.101)	(0.474)	(0.918)	(0.510)	(0.214)
CRSP Turnover	-0.215	-0.443c	-0.442c	-0.466c	-0.082	-0.263b	-0.266b	-0.310c	-0.270	-0.441b	-0.495c	-0.544c
	(0.257)	(0.003)	(0.000)	(0.000)	(0.579)	(0.035)	(0.013)	(0.002)	(0.252)	(0.019)	(0.001)	(0.000)
Idiosyncratic Volatility	-0.430	-0.615	-0.620	-0.566a	-0.661a	-0.811c	-0.827c	-0.742c	-0.523	-0.861c	-0.966c	-0.943c
	(0.495)	(0.200)	(0.131)	(0.094)	(0.089)	(0.002)	(0.000)	(0.000)	(0.104)	(0.002)	(0.000)	(0.000)
Market Beta	-0.383	-0.325	-0.274a	-0.290a	-0.161	-0.146	-0.130	-0.146	-0.346c	-0.397c	-0.300c	-0.286c
	(0.264)	(0.162)	(0.096)	(0.061)	(0.443)	(0.344)	(0.214)	(0.138)	(0.003)	(0.000)	(0.001)	(0.000)
NOBS	201,235	197,921	194,636	191,422	201,110	197,676	194,276	190,944	204,070	200,913	197,688	194,504
R^2	0.042	0.051	0.052	0.055	0.007	0.013	0.020	0.026	0.008	0.017	0.023	0.029

Panel B			EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
			Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$\boldsymbol{p}$	$\boldsymbol{\sigma(\theta)}$	$\boldsymbol{\theta}$	1.360‡	1. 232‡	1.140‡	1.018‡	1.333‡	1.190‡	1.083‡	0.948‡	0.895‡	0.851‡	0.844‡	0.766‡
0.05		0.053	0.006	0.005	0.005	0.004	0.005	0.005	0.004	0.004	0.004	0.004	0.003	0.003
0.10		0.111	0.012	0.011	0.010	0.009	0.012	0.010	0.009	0.008	0.008	0.007	0.007	0.007
0.20		0.250	0.027	0.024	0.022	0.020	0.026	0.023	0.021	0.019	0.018	0.017	0.017	0.015
0.50		1.000	0.107	0.097	0.089	0.080	0.104	0.093	0.085	0.074	0.070	0.067	0.066	0.060
0.80		4.000	0.426	0.386	0.357	0.319	0.418	0.373	0.339	0.297	0.281	0.267	0.265	0.240
0.90		9.000	0.959	0.869	0.804	0.718	0.940	0.839	0.764	0.669	0.631	0.600	0.595	0.540
0.95		19.000	2.025	1.834	1.697	1.516	1.985	1.772	1.613	1.412	1.333	1.267	1.257	1.141
	12.760	Hi-Lo	2.019	1.829	1.693	1.512	1.979	1.767	1.608	1.408	1.329	1.264	1.253	1.137
$\boldsymbol{p}$	$\boldsymbol{\sigma(\theta^2)}$	$\boldsymbol{\theta^2}$	-1.088‡	-1.031‡	-0.972‡	-0.876‡	-1.063‡	-0.984‡	-0.919‡	-0.820‡	-0.700‡	-0.706‡	-0.729‡	-0.651‡
0.05		0.003	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
0.10		0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.20		0.063	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.50		1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.80		16.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
0.90		81.000	-0.016	-0.015	-0.014	-0.013	-0.016	-0.015	-0.014	-0.012	-0.010	-0.010	-0.011	-0.010
0.95		361.00	-0.082	-0.078	-0.073	-0.066	-0.080	-0.074	-0.069	-0.062	-0.053	-0.053	-0.055	-0.049
	1,076.51	Average	-0.365	-0.346	-0.326	-0.294	-0.356	-0.330	-0.308	-0.275	-0.235	-0.237	-0.244	-0.218

Panel C	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ \equiv \theta_i^s$	1.198[c]	1.066[c]	0.977[c]	0.859[c]	1.225[c]	1.089[c]	0.990[c]	0.861[c]	0.992[c]	0.858[c]	0.825[c]	0.722[c]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ^2 \equiv \theta_i^{s\,2}$	-0.909[c]	-0.881[c]	-0.829[c]	-0.738[c]	-0.965[c]	-0.915[c]	-0.858[c]	-0.759[c]	-0.776[c]	-0.719[c]	-0.725[c]	-0.623[c]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\Delta Breadth$	0.506[c]	0.282[c]	0.122	0.062	0.430[c]	0.209[a]	0.063	0.018	0.352[c]	0.223[b]	0.055	0.007
	(0.001)	(0.010)	(0.172)	(0.395)	(0.002)	(0.053)	(0.449)	(0.782)	(0.004)	(0.015)	(0.449)	(0.917)
$\Delta Active\ MF\ Ownership$	-0.148[b]	-0.154[c]	-0.085[a]	-0.076	-0.154[b]	-0.156[c]	-0.113[b]	-0.105[b]	-0.169[b]	-0.155[c]	-0.067	-0.047
	(0.049)	(0.005)	(0.073)	(0.103)	(0.039)	(0.005)	(0.010)	(0.019)	(0.038)	(0.005)	(0.127)	(0.304)
Holding Dispersion Index	-0.036	-0.081[b]	-0.070[a]	-0.039	-0.069	-0.083[a]	-0.068[a]	-0.038	-0.013	-0.072[a]	-0.054	-0.011
	(0.450)	(0.045)	(0.059)	(0.205)	(0.191)	(0.055)	(0.068)	(0.213)	(0.790)	(0.061)	(0.110)	(0.714)
Ln(MCAP)	-0.681[c]	-0.334[a]	-0.194	-0.079	-0.282[c]	0.136[a]	0.253[c]	0.331[c]	-0.618[c]	-0.339[a]	-0.210	-0.117
	(0.004)	(0.073)	(0.227)	(0.584)	(0.006)	(0.098)	(0.001)	(0.000)	(0.008)	(0.079)	(0.193)	(0.446)
Ln(Book-to-Market)	0.113	0.181	0.203	0.238	-0.123	-0.005	0.061	0.106	0.062	0.136	0.188	0.228[b]
	(0.617)	(0.363)	(0.245)	(0.114)	(0.289)	(0.967)	(0.525)	(0.198)	(0.644)	(0.289)	(0.132)	(0.045)
Prior Year Return	-0.102	-0.096	-0.112	-0.115	-0.098	-0.073	-0.075	-0.053	0.090	0.024	-0.014	-0.018
	(0.717)	(0.688)	(0.576)	(0.510)	(0.598)	(0.558)	(0.471)	(0.581)	(0.441)	(0.805)	(0.885)	(0.855)
CRSP Turnover	-0.147	-0.407[b]	-0.420[c]	-0.465[c]	-0.094	-0.320[b]	-0.327[c]	-0.372[c]	-0.177	-0.334[b]	-0.404[c]	-0.439[c]
	(0.483)	(0.016)	(0.002)	(0.000)	(0.586)	(0.017)	(0.003)	(0.000)	(0.284)	(0.023)	(0.001)	(0.000)
Idiosyncratic Volatility	-0.885[b]	-1.114[c]	-1.115[c]	-1.034[c]	-0.771[c]	-0.949[c]	-0.961[c]	-0.877[c]	-0.837[c]	-1.172[c]	-1.241[c]	-1.251[c]
	(0.014)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Market Beta	0.015	-0.031	-0.072	-0.141	0.060	0.029	-0.004	-0.056	-0.553[c]	-0.577[c]	-0.490[c]	-0.483[c]
	(0.953)	(0.873)	(0.627)	(0.299)	(0.766)	(0.853)	(0.973)	(0.605)	(0.001)	(0.000)	(0.000)	(0.000)
Constant	0.382	-0.640	-1.036[b]	-1.225[c]	-0.418	-1.406[c]	-1.693[c]	-1.807[c]	0.442[b]	-0.541[c]	-0.909[c]	-1.122[c]
	(0.504)	(0.196)	(0.014)	(0.001)	(0.204)	(0.000)	(0.000)	(0.000)	(0.019)	(0.009)	(0.000)	(0.000)
NOBS	201,235	197,921	194,636	191,422	201,110	197,676	194,276	190,944	204,070	200,913	197,688	194,504
R^2	0.067	0.073	0.077	0.080	0.031	0.036	0.038	0.041	0.029	0.038	0.045	0.052

Table VIII: Forecast Returns: Fund Performance using Management Fees

Table reports style and quarter fixed effects regressions of lead quarter buy-and-hold stock returns on investment quality and control variables. ‡indicates that management fees are used to proxy fund performance. The cross-product of active holding and management fees summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. Average quarterly returns are expressed in percent. Variable definitions can be found in Table II. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. In two-way fixed effects regressions, errors are clustered by style and quarter. p-values are reported in parentheses. Superscript $^{a, b, c}$ denote statistical significance at 10%, 5%, and 1% levels.

	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ^{\ddagger} \equiv \theta_i^s$	1.328^{c}	1.283^{c}	1.145^{c}	1.077^{c}	1.364^{c}	1.273^{c}	1.110^{c}	1.022^{c}	0.881^{c}	0.960^{c}	0.912^{c}	0.881^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ^2 \equiv \theta_i^{s\,2}$	-1.106^{c}	-1.107^{c}	-1.001^{c}	-0.971^{c}	-1.136^{c}	-1.088^{c}	-0.961^{c}	-0.916^{c}	-0.747^{c}	-0.856^{c}	-0.828^{c}	-0.803^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\Delta Breadth$	0.529^{c}	0.243^{b}	0.096	0.024	0.492^{c}	0.236^{b}	0.107	0.042	0.309^{c}	0.228^{c}	0.060	0.011
	(0.005)	(0.032)	(0.273)	(0.748)	(0.001)	(0.011)	(0.128)	(0.492)	(0.009)	(0.006)	(0.357)	(0.845)
$\Delta Active\ MF\ Ownership$	-0.060	-0.110^{a}	-0.069	-0.043	-0.073	-0.097^{a}	-0.070	-0.053	-0.043	-0.098^{a}	-0.034	-0.015
	(0.385)	(0.064)	(0.241)	(0.382)	(0.275)	(0.059)	(0.184)	(0.235)	(0.520)	(0.057)	(0.510)	(0.731)
Holding Dispersion Index	-0.002	-0.054	-0.053	-0.025	-0.047	-0.068^{a}	-0.059^{a}	-0.030	-0.000	-0.068^{b}	-0.050	-0.009
	(0.970)	(0.141)	(0.120)	(0.375)	(0.341)	(0.073)	(0.079)	(0.278)	(0.993)	(0.045)	(0.110)	(0.736)
Ln(MCAP)	-0.814	-0.388	-0.282	-0.206	-0.988^{a}	-0.454	-0.211	-0.057	-0.658	-0.437	-0.439	-0.400
	(0.138)	(0.343)	(0.388)	(0.450)	(0.055)	(0.202)	(0.443)	(0.789)	(0.330)	(0.434)	(0.328)	(0.323)
Ln(Book-to-Market)	0.670	0.825^{b}	0.792^{c}	0.791^{c}	0.110	0.285	0.330^{a}	0.371^{b}	-0.032	0.230	0.358^{b}	0.433^{c}
	(0.179)	(0.015)	(0.002)	(0.001)	(0.741)	(0.198)	(0.058)	(0.014)	(0.889)	(0.182)	(0.021)	(0.002)
Prior Year Return	-0.592	-0.542	-0.513^{a}	-0.501^{a}	-0.315	-0.247	-0.243	-0.237^{a}	0.116	-0.012	-0.067	-0.114
	(0.217)	(0.166)	(0.098)	(0.058)	(0.306)	(0.233)	(0.152)	(0.099)	(0.480)	(0.921)	(0.505)	(0.215)
CRSP Turnover	-0.225	-0.456^{c}	-0.452^{c}	-0.478^{c}	-0.095	-0.278^{b}	-0.276^{b}	-0.321^{c}	-0.278	-0.453^{b}	-0.504^{c}	-0.555^{c}
	(0.233)	(0.003)	(0.000)	(0.000)	(0.522)	(0.026)	(0.010)	(0.002)	(0.240)	(0.016)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.417	-0.608	-0.615	-0.563^{a}	-0.650^{a}	-0.804^{c}	-0.823^{c}	-0.740^{c}	-0.513	-0.856^{c}	-0.963^{c}	-0.939^{c}
	(0.509)	(0.206)	(0.135)	(0.096)	(0.095)	(0.003)	(0.000)	(0.000)	(0.113)	(0.002)	(0.000)	(0.000)
Market Beta	-0.386	-0.327	-0.275^{a}	-0.292^{a}	-0.164	-0.148	-0.131	-0.148	-0.350^{c}	-0.400^{c}	-0.301^{c}	-0.288^{c}
	(0.261)	(0.160)	(0.095)	(0.059)	(0.436)	(0.339)	(0.210)	(0.133)	(0.002)	(0.000)	(0.001)	(0.000)
NOBS	201,234	197,922	194,637	191,422	201,110	197,677	194,277	190,944	204,070	200,913	197,689	194,505
R^2	0.042	0.051	0.052	0.055	0.007	0.013	0.020	0.026	0.008	0.017	0.023	0.029

Table IX: Forecast Returns: Investment Quality of Stocks by Market Capitalization

Table Panel A reports two-way fixed effects regressions of lead quarter buy-and-hold stock returns on the investment quality of small cap, midcap and large cap stocks, as well as control variables. Panel B reports forecasted returns at selected percentiles of selection quality between the 5th and the 95th percentiles for each tercile of market capitalization.‡ indicates estimated coefficients from Panel A. Stocks are sorted into terciles by market capitalization using NYSE stocks to establish breakpoints. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. Average quarterly returns are expressed in percent. Variable definitions can be found in Table II. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. Errors are clustered by style and quarter. p-values are reported in parentheses. Superscript a, b, c denote statistical significance at 10%, 5%, and 1% levels. Unreported estimated coefficients on control variables are reported in an Appendix Table IV.

Panel A	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ \times Small\ Cap$	1.569^{c}	1.447^{c}	1.304^{c}	1.135^{c}	1.557^{c}	1.385^{c}	1.227^{c}	1.050^{c}	1.020^{c}	0.968^{c}	0.960^{c}	0.844^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ \times Mid\ Cap$	1.306^{c}	1.089^{c}	1.007^{c}	0.911^{c}	1.304^{c}	1.081^{c}	0.972^{c}	0.879^{c}	0.890^{c}	0.789^{c}	0.784^{c}	0.733^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ \times Large\ Cap$	1.114^{c}	1.076^{c}	1.053^{c}	0.971^{c}	1.037^{c}	1.028^{c}	0.999^{c}	0.881^{c}	0.718^{c}	0.749^{c}	0.743^{c}	0.690^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ^2 \equiv \theta_i^{s2}$	-1.241^{c}	-1.185^{c}	-1.088^{c}	-0.958^{c}	-1.228^{c}	-1.124^{c}	-1.021^{c}	-0.892^{c}	-0.791^{c}	-0.789^{c}	-0.812^{c}	-0.708^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\Delta Breadth$	0.533^{c}	0.248^{b}	0.100	0.028	0.497^{c}	0.241^{c}	0.111	0.046	0.312^{c}	0.231^{c}	0.063	0.015
	(0.005)	(0.029)	(0.252)	(0.704)	(0.001)	(0.010)	(0.113)	(0.449)	(0.009)	(0.006)	(0.334)	(0.800)
$\Delta Active\ MF\ Ownership$	-0.062	-0.112^{a}	-0.071	-0.044	-0.075	-0.099^{a}	-0.071	-0.054	-0.044	-0.099^{a}	-0.035	-0.016
	(0.373)	(0.061)	(0.231)	(0.367)	(0.264)	(0.055)	(0.175)	(0.223)	(0.513)	(0.054)	(0.496)	(0.711)
Holding Dispersion Index	-0.006	-0.058	-0.056	-0.028	-0.051	-0.072^{a}	-0.062^{a}	-0.033	-0.004	-0.072^{b}	-0.052^{a}	-0.011
	(0.902)	(0.115)	(0.100)	(0.317)	(0.295)	(0.056)	(0.064)	(0.230)	(0.931)	(0.034)	(0.090)	(0.661)
NOBS	201,235	197,921	194,636	191,422	201,110	197,676	194,276	190,944	204,070	200,913	197,688	194,504
R^2	0.042	0.051	0.052	0.055	0.007	0.013	0.020	0.026	0.008	0.017	0.023	0.029
$Large/Small\ Cap$	71.00%	74.36%	80.75%	85.55%	66.60%	74.22%	81.42%	83.90%	70.39%	77.38%	77.40%	81.75%
$Mid\ Cap/Small\ Cap$	83.24%	75.26%	77.22%	80.26%	83.75%	78.05%	79.22%	83.71%	87.25%	81.51%	81.67%	86.85%

Panel B			EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
p	$\sigma(\theta)$	θ	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
Small Cap			1.569‡	1.447‡	1.304‡	1.135‡	1.557‡	1.385‡	1.227‡	1.050‡	1.020‡	0.968‡	0.960‡	0.844‡
0.05		0.053	0.006	0.006	0.005	0.005	0.006	0.006	0.005	0.004	0.004	0.004	0.004	0.003
0.10		0.111	0.014	0.013	0.011	0.010	0.014	0.012	0.011	0.009	0.009	0.008	0.008	0.007
0.20		0.250	0.031	0.028	0.026	0.022	0.031	0.027	0.024	0.021	0.020	0.019	0.019	0.017
0.50		1.000	0.123	0.113	0.102	0.089	0.122	0.109	0.096	0.082	0.080	0.076	0.075	0.066
0.80		4.000	0.492	0.454	0.409	0.356	0.488	0.434	0.385	0.329	0.320	0.303	0.301	0.265
0.90		9.000	1.107	1.021	0.920	0.801	1.098	0.977	0.865	0.741	0.719	0.683	0.677	0.595
0.95		19.000	2.336	2.155	1.942	1.690	2.318	2.062	1.827	1.563	1.519	1.441	1.429	1.257
	12.760	Hi-Lo	2.330	2.149	1.936	1.685	2.312	2.057	1.822	1.559	1.515	1.437	1.425	1.253
Mid Cap			1.306‡	1.089‡	1.007‡	0.911‡	1.304‡	1.081‡	0.972‡	0.879‡	0.890‡	0.789‡	0.784‡	0.733‡
0.05		0.053	0.005	0.004	0.004	0.004	0.005	0.004	0.004	0.004	0.004	0.003	0.003	0.003
0.10		0.111	0.011	0.009	0.009	0.008	0.011	0.009	0.008	0.008	0.008	0.007	0.007	0.006
0.20		0.250	0.026	0.021	0.020	0.018	0.026	0.021	0.019	0.017	0.017	0.015	0.015	0.014
0.50		1.000	0.102	0.085	0.079	0.071	0.102	0.085	0.076	0.069	0.070	0.062	0.061	0.057
0.80		4.000	0.409	0.341	0.316	0.286	0.409	0.339	0.305	0.276	0.279	0.247	0.246	0.230
0.90		9.000	0.921	0.768	0.710	0.643	0.920	0.762	0.686	0.620	0.628	0.556	0.553	0.517
0.95		19.000	1.945	1.622	1.499	1.356	1.942	1.610	1.447	1.309	1.325	1.175	1.167	1.091
	12.760	Hi-Lo	1.939	1.617	1.495	1.353	1.936	1.605	1.443	1.305	1.322	1.172	1.164	1.088
Large Cap			1.114‡	1.076‡	1.053‡	0.971‡	1.037‡	1.028‡	0.999‡	0.881‡	0.718‡	0.749‡	0.743‡	0.690‡
0.05		0.053	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.003	0.003	0.003	0.003
0.10		0.111	0.010	0.009	0.009	0.008	0.009	0.009	0.009	0.008	0.006	0.007	0.006	0.006
0.20		0.250	0.022	0.021	0.021	0.019	0.020	0.020	0.020	0.017	0.014	0.015	0.015	0.014
0.50		1.000	0.087	0.084	0.083	0.076	0.081	0.081	0.078	0.069	0.056	0.059	0.058	0.054
0.80		4.000	0.349	0.337	0.330	0.304	0.325	0.322	0.313	0.276	0.225	0.235	0.233	0.216
0.90		9.000	0.786	0.759	0.743	0.685	0.731	0.725	0.705	0.621	0.506	0.528	0.524	0.487
0.95		19.000	1.659	1.602	1.568	1.446	1.544	1.531	1.488	1.312	1.069	1.115	1.106	1.027
	12.760	Hi-Lo	1.654	1.598	1.564	1.442	1.540	1.526	1.483	1.308	1.066	1.112	1.103	1.025

Table X: Forecast Returns: Investment Quality and Conviction Quality of Stocks

Table reports two-way fixed effects regressions of lead quarter buy-and-hold stock returns on investment quality and conviction quality, as well as control variables. Panel A presents the average effect of investment quality and conviction quality. Panel B presents the interaction of investment quality and conviction quality. Panel C presents summary statistics on conviction quality in the selected percentiles between the 5th and the 95th percentiles. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. The cross-product of active holding turnover and GVA summed across funds is used to identify active turnover – the patience and conviction of fund managers. Conviction quality $Turn \equiv \theta_i^\pi$ is the odds ratio of a stock's relative percentile rank, θ_i^π, on active turnover. Each quarter, Hi_IQ equals 1 when investment quality is above median and 0 otherwise, Lo_IQ equals 1 when investment quality is below median and 0 otherwise. Interaction terms of Hi_IQ with conviction quality and Lo_IQ with conviction quality are used as regressors in Panel B. Average quarterly returns are expressed in percent. Variable definitions can be found in Table II. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. Errors are clustered by style and quarter. p-values are reported in parentheses. Superscripts $^{a, b, c}$ denote statistical significance at 10%, 5%, and 1% levels. Unreported estimated coefficients on control variables are reported in Appendix Table IV.

	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
Panel A	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ \equiv \theta_i^s$	1.320^{c}	1.202^{c}	1.092^{c}	0.983^{c}	1.282^{c}	1.128^{c}	1.022^{c}	0.900^{c}	0.888^{c}	0.841^{c}	0.795^{c}	0.731^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Turn \equiv \theta_i^\pi$	-1.007^{c}	-0.939^{c}	-0.861^{c}	-0.797^{c}	-0.970^{c}	-0.875^{c}	-0.809^{c}	-0.735^{c}	-0.684^{c}	-0.669^{c}	-0.647^{c}	-0.603^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ^2 \equiv \theta_i^{s^2}$	-0.407^{c}	-0.307^{c}	-0.229^{c}	-0.199^{c}	-0.348^{c}	-0.239^{c}	-0.171^{c}	-0.151^{c}	-0.199^{c}	-0.127^{b}	-0.085	-0.067
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.026)	(0.116)	(0.171)
$\Delta Breadth$	0.361^{b}	0.185^{a}	0.047	-0.020	0.379^{c}	0.199^{b}	0.089	0.016	0.269^{b}	0.217^{b}	0.060	0.002
	(0.019)	(0.082)	(0.564)	(0.769)	(0.004)	(0.034)	(0.199)	(0.783)	(0.017)	(0.011)	(0.334)	(0.973)
$\Delta Active\ MF\ Ownership$	-0.039	-0.094	-0.047	-0.018	-0.061	-0.079	-0.051	-0.029	-0.041	-0.095^{a}	-0.027	-0.003
	(0.540)	(0.112)	(0.413)	(0.706)	(0.367)	(0.106)	(0.295)	(0.513)	(0.557)	(0.083)	(0.609)	(0.939)
Holding Dispersion Index	0.024	-0.047	-0.039	-0.012	-0.021	-0.052	-0.042	-0.014	0.036	-0.054	-0.037	0.004
	(0.605)	(0.192)	(0.251)	(0.652)	(0.629)	(0.138)	(0.176)	(0.568)	(0.400)	(0.108)	(0.209)	(0.883)
NOBS	181,967	179,045	176,154	173,311	181,852	178,822	175,808	172,855	184,465	181,672	178,830	176,027
R^2	0.036	0.043	0.044	0.046	0.005	0.012	0.019	0.024	0.008	0.016	0.021	0.027
Relative to Lead 1 Qtr	1.000	0.911	0.827	0.745	1.000	0.880	0.797	0.702	1.000	0.947	0.895	0.823

Panel B	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$Hi\_IQ \times Turn$	-0.421[c]	-0.321[c]	-0.256[c]	-0.227[c]	-0.359[c]	-0.243[c]	-0.187[c]	-0.169[c]	-0.183[c]	-0.117[a]	-0.086	-0.075
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.051)	(0.137)	(0.137)
$Lo\_IQ \times Turn$	-0.207	-0.115	0.021	0.046	-0.165	-0.109	0.015	0.032	-0.202	-0.097	-0.007	0.043
	(0.288)	(0.398)	(0.862)	(0.672)	(0.362)	(0.384)	(0.890)	(0.748)	(0.209)	(0.512)	(0.956)	(0.738)
$IQ^2 \equiv \theta_i^{s^2}$	0.200[b]	0.160[b]	0.123[b]	0.086[a]	0.205[c]	0.168[c]	0.122[c]	0.082[a]	0.155[b]	0.120[a]	0.090	0.067
	(0.011)	(0.017)	(0.019)	(0.094)	(0.005)	(0.009)	(0.010)	(0.087)	(0.034)	(0.056)	(0.106)	(0.212)
$\Delta Breadth$	0.363[b]	0.186[a]	0.048	-0.019	0.381[c]	0.200[b]	0.089	0.017	0.271[b]	0.218[b]	0.060	0.002
	(0.019)	(0.081)	(0.561)	(0.774)	(0.004)	(0.033)	(0.196)	(0.777)	(0.017)	(0.010)	(0.330)	(0.966)
$\Delta Active\ MF\ Ownership$	-0.036	-0.091	-0.044	-0.016	-0.057	-0.076	-0.049	-0.027	-0.039	-0.093[a]	-0.025	-0.002
	(0.573)	(0.125)	(0.443)	(0.745)	(0.390)	(0.119)	(0.320)	(0.546)	(0.578)	(0.090)	(0.635)	(0.971)
Holding Dispersion Index	0.019	-0.051	-0.043	-0.016	-0.026	-0.055	-0.045	-0.017	0.034	-0.056[a]	-0.039	0.001
	(0.681)	(0.160)	(0.210)	(0.557)	(0.567)	(0.117)	(0.148)	(0.486)	(0.433)	(0.097)	(0.185)	(0.960)
NOBS	181,967	179,045	176,154	173,311	181,852	178,822	175,808	172,855	184,465	181,672	178,830	176,027
R^2	0.036	0.042	0.043	0.045	0.005	0.012	0.018	0.023	0.008	0.015	0.021	0.027

Panel C

Active Turnover Percentile Rank p	Change in Holdings (ΔAH)	GVA	$\rho\left(\begin{matrix}\Delta AH, \\ GVA\end{matrix}\right)$	MCAP	Book-to-Market
0.05	0.016	5.40	0.008	5,136	0.850
0.10	0.031	5.86	-0.002	4,900	0.731
0.20	0.036	5.91	-0.015	7,326	0.699
0.50	0.031	5.38	-0.021	3,811	0.703
0.80	0.034	5.91	-0.003	6,721	0.683
0.90	0.030	5.74	0.011	5,587	0.737
0.95	0.016	5.38	0.027	4,245	0.814

Table XI: Forecast Returns: Investment Quality and Conviction Quality of Stocks by Market Capitalization

Table reports two-way fixed effects regressions of lead quarter buy-and-hold stock returns on the investment quality and conviction quality of small cap, midcap and large cap stocks, as well as control variables. Average quarterly returns are expressed in percent. Stocks are sorted into terciles by market capitalization using NYSE stocks to establish breakpoint. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. The cross-product of active holding turnover and GVA summed across funds is used to identify active turnover – the patience and conviction of fund managers. Conviction quality $Turn \equiv \theta_i^\pi$ is the odds ratio of a stock's relative percentile rank, θ_i^π, on active turnover. Variable definitions can be found in Table II. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. Errors are clustered by style and quarter. p-values are reported in parentheses. Superscript $^{a, b, c}$ denote statistical significance at 10%, 5%, and 1% levels. Unreported estimated coefficients on control variables are reported in Appendix Table IV.

	EXCESS MARKET RETURN				**DGTW RETURN**				**4-FACTOR ALPHA**			
	Lead 1 Qtr	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**	**Lead 1 Qtr**	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**	**Lead 1 Qtr**	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**
$IQ \times Small\ Cap$	1.422^{c}	1.342^{c}	1.215^{c}	1.085^{c}	1.408^{c}	1.263^{c}	1.136^{c}	0.994^{c}	0.939^{c}	0.921^{c}	0.903^{c}	0.829^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ \times Midcap$	1.297^{c}	1.105^{c}	1.079^{c}	0.993^{c}	1.353^{c}	1.093^{c}	1.039^{c}	0.941^{c}	1.075^{c}	0.967^{c}	0.957^{c}	0.903^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ \times Large\ Cap$	1.283^{c}	1.289^{c}	1.213^{c}	1.100^{c}	1.178^{c}	1.141^{c}	1.090^{c}	0.963^{c}	0.833^{c}	0.875^{c}	0.813^{c}	0.755^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
IQ^2	-1.083^{c}	-1.043^{c}	-0.956^{c}	-0.876^{c}	-1.067^{c}	-0.975^{c}	-0.897^{c}	-0.809^{c}	-0.729^{c}	-0.735^{c}	-0.735^{c}	-0.686^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Turn \times Small\ Cap$	-0.400^{c}	-0.292^{c}	-0.208^{c}	-0.179^{c}	-0.336^{c}	-0.226^{c}	-0.153^{c}	-0.134^{c}	-0.184^{c}	-0.107^{a}	-0.062	-0.043
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.002)	(0.002)	(0.064)	(0.258)	(0.378)
$Turn \times Midcap$	-0.475	-0.332	-0.637^{a}	-0.673^{b}	-0.615	-0.397	-0.592	-0.602^{a}	-0.738	-0.750^{a}	-0.953^{b}	-0.991^{c}
	(0.326)	(0.362)	(0.082)	(0.047)	(0.189)	(0.249)	(0.105)	(0.068)	(0.156)	(0.055)	(0.014)	(0.009)
$Turn \times Large\ Cap$	-0.720	-1.133^{c}	-1.146^{c}	-1.007^{c}	-0.790^{c}	-0.848^{b}	-0.912^{c}	-0.824^{c}	-0.745^{b}	-0.889^{c}	-0.882^{c}	-0.813^{c}
	(0.174)	(0.003)	(0.000)	(0.000)	(0.003)	(0.011)	(0.001)	(0.001)	(0.011)	(0.008)	(0.002)	(0.002)
$\Delta Breadth$	0.361^{b}	0.186^{a}	0.048	-0.019	0.379^{c}	0.199^{b}	0.089	0.017	0.269^{b}	0.217^{b}	0.060	0.002
	(0.020)	(0.082)	(0.560)	(0.774)	(0.004)	(0.034)	(0.196)	(0.778)	(0.017)	(0.011)	(0.332)	(0.968)
$\Delta Active\ MF\ Ownership$	-0.039	-0.094	-0.047	-0.019	-0.061	-0.080	-0.052	-0.029	-0.042	-0.096^{a}	-0.028	-0.004
	(0.537)	(0.109)	(0.405)	(0.695)	(0.363)	(0.104)	(0.289)	(0.503)	(0.550)	(0.080)	(0.596)	(0.923)
Holding Dispersion Index	0.024	-0.046	-0.038	-0.011	-0.021	-0.051	-0.041	-0.013	0.037	-0.053	-0.035	0.005
	(0.601)	(0.198)	(0.264)	(0.678)	(0.638)	(0.143)	(0.186)	(0.592)	(0.390)	(0.117)	(0.228)	(0.843)
NOBS	181,967	179,045	176,154	173,311	181,852	178,822	175,808	172,855	184,465	181,672	178,830	176,027
R^2	0.036	0.043	0.044	0.046	0.005	0.013	0.019	0.024	0.008	0.016	0.022	0.027

Table XII: Forecast Returns: Investment Quality and Conviction Quality predicted from funds in GVA terciles

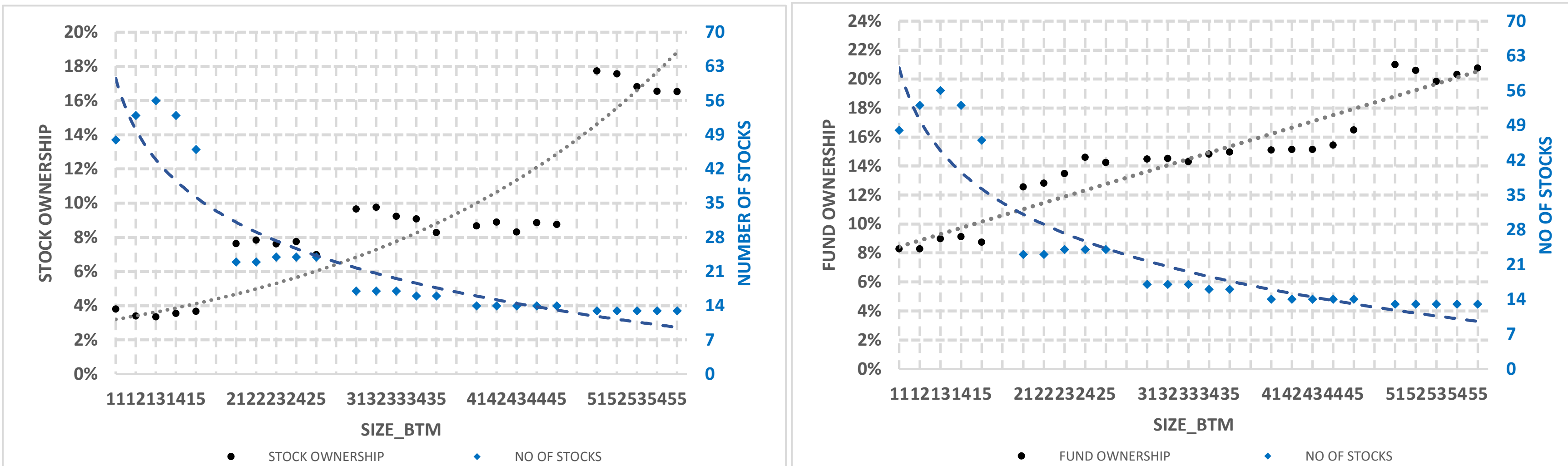
Table reports two-way fixed effects regressions of lead quarter buy-and-hold stock returns on the investment quality and conviction quality estimated from three fund GVA terciles, as well as control variables. At the end of each quarter, funds are sorted into terciles by their GVAs. The cross-products of active holding and GVA summed across funds in each tercile are used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. The cross-products of active holding turnover and GVA summed across funds in each tercile are used to identify active turnover – the patience and conviction of fund managers. Conviction quality $Turn \equiv \theta_i^\pi$ is the odds ratio of a stock's relative percentile rank, θ_i^π, on active turnover. Variable definitions can be found in Table II. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. Errors are clustered by style and quarter. Average quarterly returns are expressed in percent. p-values are reported in parentheses. Superscript $^{a,\,b,\,c}$ denote statistical significance at 10%, 5%, and 1% levels. Unreported estimated coefficients on control variables are reported in an Appendix Table IV.

	EXCESS MARKET RETURN				**DGTW RETURN**				**4-FACTOR ALPHA**			
	Lead 1 Qtr	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**	**Lead 1 Qtr**	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**	**Lead 1 Qtr**	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**
IQ: top GVA tercile	0.356^{c}	0.334^{c}	0.298^{c}	0.245^{c}	0.376^{c}	0.336^{c}	0.290^{c}	0.238^{c}	0.243^{c}	0.241^{c}	0.220^{c}	0.197^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.001)
IQ: middle GVA tercile	0.074	0.064	0.092^{b}	0.101^{c}	0.109^{a}	0.088^{a}	0.094^{b}	0.102^{c}	0.078	0.085	0.099^{b}	0.108^{b}
	(0.276)	(0.193)	(0.045)	(0.008)	(0.092)	(0.054)	(0.032)	(0.005)	(0.230)	(0.102)	(0.041)	(0.012)
IQ: bottom GVA tercile	-0.085	-0.091	-0.058	-0.014	-0.077	-0.023	-0.004	0.032	0.142^{a}	0.134^{a}	0.154^{b}	0.157^{c}
	(0.474)	(0.296)	(0.442)	(0.811)	(0.476)	(0.741)	(0.943)	(0.529)	(0.077)	(0.065)	(0.025)	(0.005)
$Turn$: top GVA tercile	-0.312^{c}	-0.298^{c}	-0.239^{c}	-0.217^{c}	-0.287^{c}	-0.285^{c}	-0.227^{c}	-0.215^{c}	-0.209^{c}	-0.207^{c}	-0.197^{b}	-0.184^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)	(0.012)	(0.004)
$Turn$: middle GVA	-0.044	0.051	0.049	0.047	-0.015	0.088^{a}	0.085	0.088	-0.022	0.068	0.095	0.098
	(0.529)	(0.396)	(0.498)	(0.445)	(0.817)	(0.100)	(0.174)	(0.109)	(0.807)	(0.270)	(0.198)	(0.132)
$Turn$: bottom GVA	-0.185^{a}	-0.115	-0.098	-0.109	-0.132	-0.065	-0.033	-0.034	-0.030	0.050	0.058	0.042
	(0.094)	(0.190)	(0.189)	(0.124)	(0.122)	(0.306)	(0.558)	(0.544)	(0.663)	(0.423)	(0.301)	(0.419)
IQ^2: top GVA tercile	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000
	(0.853)	(0.239)	(0.195)	(0.352)	(0.853)	(0.282)	(0.135)	(0.233)	(0.701)	(0.400)	(0.277)	(0.411)
IQ^2: middle GVA tercile	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	0.000	0.000	-0.000	-0.000
	(0.532)	(0.453)	(0.392)	(0.193)	(0.559)	(0.772)	(0.953)	(0.701)	(0.847)	(0.947)	(0.473)	(0.282)
IQ^2: bottom GVA tercile	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.166)	(0.153)	(0.516)	(0.269)	(0.236)	(0.221)	(0.753)	(0.501)	(0.410)	(0.167)	(0.185)	(0.116)
NOBS	181,785	178,871	175,989	173,153	181,671	178,649	175,644	172,697	184,268	181,490	178,656	175,862
R^2	0.036	0.042	0.043	0.045	0.005	0.012	0.018	0.023	0.008	0.016	0.021	0.027

Table XIII: Cumulative Abnormal Returns around Earnings Announcements

Table reports two-way fixed effects regressions of cumulative abnormal returns (CAR1) in the three-day window [-1, 1] around earnings announcement dates each quarter, and cumulative abnormal returns (CAR2) beginning from the third day post-earnings announcement date ending on the 60th day post-earnings announcement date, or ending on the earnings announcement date in the subsequent quarter, whichever comes earlier. Earnings data and earnings announcement dates are from the IBES database. At June end of each year t, stocks are sorted into 2×3 benchmark portfolios by size (ME) and book-to-market equity (BE/ME). Median ME on NYSE stocks and 30th and 70th percentiles of BE/ME on NYSE stocks, computed as book equity in the last fiscal year end in $t-1$ divided by ME in December of $t-1$, are used as breakpoints. Abnormal returns are computed as daily returns in excess of the benchmark to which the stock belongs. Daily portfolio returns are value-weighted daily abnormal returns across stocks in the portfolio. p-values are reported in parentheses. Superscript [a, b, c] denote statistical significance at 10%, 5%, and 1% levels.

	CAR1 [-1, 1]	**CAR2 [3, 60]**			
	Lead 1 Qtr	**Lead 1 Qtr**	**Lead 2 Qtr**	**Lead 3 Qtr**	**Lead 4 Qtr**
$IQ \equiv \theta_i^s$	0.198[c]	0.502[c]	0.555[c]	0.475[b]	0.313
	(0.002)	(0.003)	(0.002)	(0.014)	(0.104)
$IQ^2 \equiv \theta_i^{s\,2}$	-0.211[c]	-0.375[b]	-0.520[c]	-0.426[b]	-0.256
	(0.002)	(0.018)	(0.001)	(0.025)	(0.197)
$\Delta Breadth$	0.148[c]	0.030	-0.096	-0.134[a]	0.050
	(0.000)	(0.818)	(0.277)	(0.080)	(0.542)
$\Delta Active\ MF\ Ownership$	-0.035	0.005	-0.058	-0.080	-0.175[b]
	(0.408)	(0.966)	(0.520)	(0.286)	(0.032)
Holding Dispersion Index	0.004	-0.089[b]	-0.004	-0.049	-0.014
	(0.813)	(0.020)	(0.934)	(0.274)	(0.755)
Ln(MCAP)	0.079	0.318	-0.215	-0.142	-0.376[a]
	(0.307)	(0.216)	(0.371)	(0.570)	(0.094)
Ln(Book-to-Market)	0.052	-0.124	-0.128	-0.125	-0.198
	(0.223)	(0.487)	(0.468)	(0.476)	(0.219)
Prior Year Return	-0.032	-0.370[a]	-0.143	-0.038	0.056
	(0.449)	(0.067)	(0.418)	(0.801)	(0.675)
CRSP Turnover	-0.181[c]	-0.168	-0.128	-0.182	-0.122
	(0.000)	(0.365)	(0.456)	(0.311)	(0.441)
Idiosyncratic Volatility	0.094	1.035[c]	0.724[c]	0.606[c]	0.613[c]
	(0.208)	(0.000)	(0.004)	(0.006)	(0.000)
Market Beta	-0.020	0.009	0.129	0.109	0.043
	(0.505)	(0.950)	(0.410)	(0.479)	(0.760)
NOBS	136,775	136,777	132,710	128,824	124,965
R^2	0.002	0.012	0.011	0.011	0.011



PANEL A PANEL B

Figure 1: Characteristics of Fund Ownership of Stocks by Style Segments

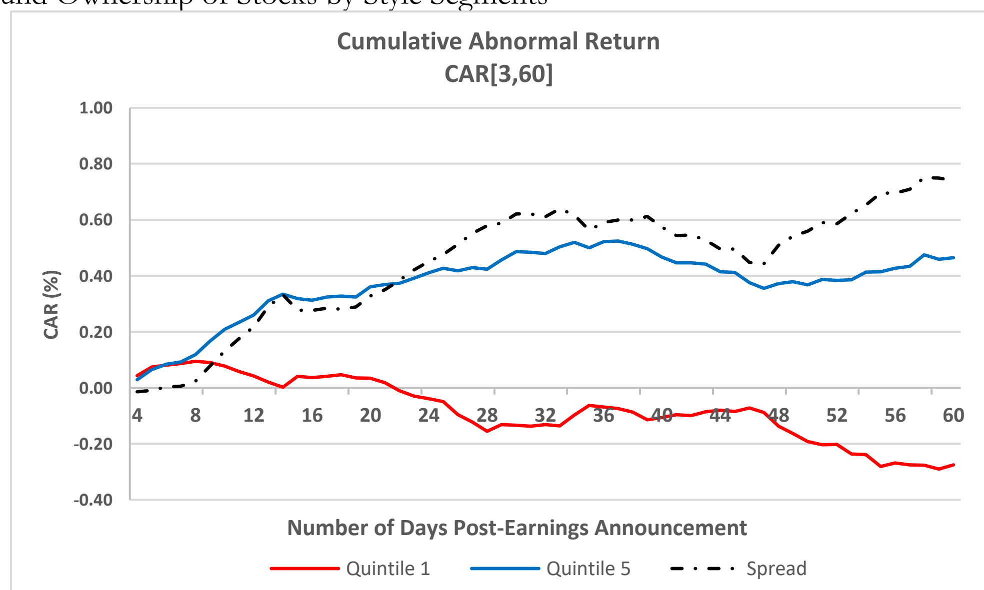


Figure 2: Cumulative Abnormal Returns on Stocks sorted into Quintiles by Stock IQ

APPENDIX

Table AI: Variable Definitions

Style Segments: Cutoffs from annual DGTW (Daniel et al. 1997) sorts in July each year of stocks into quintiles by size, industry-adjusted book-to-market, and momentum are used to assign every CRSP stock i into k = 125 style segments each quarter over our 72-quarter sample period 2000-2017.

Active Holding: Fund holding of stock i owned by fund j, $h_{i(k),j}$, is the market value of stock i owned by fund j as a percentage of all stock holdings of fund j in segment style k minus the peer group weight of stock i in segment style k.

$$h_{i(k),j} = \left(shrown_{i(k),j} \cdot prc_{i(k)}\right) / \sum_{i(k)\epsilon I(j,k)} \left(shrown_{i(k),j} \cdot prc_{i(k)}\right)$$

where $prc_{i(k)}$ and $shrown_{i(k),j}$ denote the price and shares of stock i in style segment k owned by fund j. Peer group holding, $\bar{h}_{i(k)}$, is the total market value of stock i owned by all actively managed mutual funds j as a percentage of the total market value of all stocks in segment style k owned by all actively managed mutual funds j.

$$\bar{h}_{i(k)} = \sum_{j\epsilon J(k)} \left(shrown_{i(k),j} \cdot prc_{i(k)}\right) / \sum_{j\epsilon J(k)} \sum_{i(k)\epsilon I(j,k)} \left(shrown_{i(k),j} \cdot prc_{i(k)}\right)$$

where $j\epsilon J(k)$ are the set of funds that own stocks in style segment k. Active holding, w_{ij}, is the deviation of fund from peer group holding of stock i in style segment k.

$$w_{i,j} = h_{i(k),j} - \bar{h}_{i(k)}$$

Active Holdings Turnover: Active turnover of stock i owned by fund j is the difference in active holding between the current and prior 4 quarters.

Investment Quality: For stock *i*, the sum product of active holdings and fund quality over all funds in the style segment which stock *I* belongs to. We use the odds ratio of the sum product as investment quality in the regressions.

Conviction Quality: For stock *i*, the sum product of active holdings turnover and fund quality over all funds in the style segment which stock *I* belongs to. We use the odds ratio of the sum product as conviction quality in the regressions.

Gross Alpha: Estimated from rolling 12-month time series regressions of monthly fund gross return on Fama and French (1992) market return, SMB and HML factors and Carhart(1997) momentum factor, and as in Berk and Binsbergen (2015), averaged across the current and prior months over our sample period. Monthly gross fund return is the sum of fund monthly net return and 1/12 of fund expense ratio.

TNA: Monthly total net assets under management averaged across the current and prior months over our sample period.

Gross Value-Added: Product of fund monthly gross alpha and TNA averaged across the current and prior months over our sample period. Monthly GVA is then multiplied by 3 to get quarterly fund GVA.

Management fees: Monthly fund management fee is estimated as sum of product of share class TNA at month end and monthly management fee ratio over all share classes for each fund, and then averaged across the current and prior months over our sample period. Monthly fund management fee multiplied by 3 is quarterly management fees.

Industry Concentration: In each quarter, industry concentration is computed as the squared differences between industry weights of funds, $w_{j,\mathrm{I}}$, and aggregate industry weights, w_{I}, summed across 10 broadly defined Fama and French (1997) industries.

$$Industry\ concentration_{j,\mathrm{I}} = \sum_{\mathrm{I}=1}^{10}(w_{j,\mathrm{I}} - w_{\mathrm{I}})^2$$

Dispersion Index of Active Holding: Standard deviation of active holdings across all funds j with non-zero holdings in segment style k divided by the mean active holding.

Breadth: $Ln(N)$, where N is the number of actively managed mutual funds with non-zero holdings of stock i in segment style k. Delta breadth is computed as the change in breadth from the prior quarter.

Active Mutual Fund Ownership: The percentage of total shares outstanding of stock i owned by actively managed mutual funds j at the end of quarter q. Quarterly change in active mutual fund ownership is computed as the change in active mutual fund ownership from the prior quarter.

Market Capitalization: Market equity capitalization is the product of closing price and total shares outstanding of stock i at the end of the quarter q expressed in millions of dollars. We use natural log of market capitalization in regressions.

Book-to-Market: Book equity to shareholders' equity in Daniel and Titman (2006). We use nature log of book to market ratio in regressions.

Prior year return: In month t at quarter end is the cumulative monthly return over the prior 12 months starting from $t - 2$ and ending in $t - 13$.

CRSP Turnover: Total trading volume reported by CRSP summed across all 3 months in the quarter as a percentage of total shares outstanding where trading volume is adjusted following French (2008).

Idiosyncratic Volatility: Standard deviation of residuals from time series regressions of daily stock returns on Fama French (1992) market, SMB and HML factors over the quarter.

Market Beta: Sum of the coefficients on contemporaneous and five lags of market excess returns estimated from time series regressions of daily stock excess returns on daily contemporaneous and five lags of market excess returns each quarter following Jiang and Sun (2014).

Standardized Earnings Surprise (SUE): For each stock in the quarter, we compute the SUE as actual earnings per share minus median analyst forecasts made earlier than earnings announcement date but no more than 90 days in advance expressed as a percent of stock price at the end of quarter. If there are multiple forecasts from the same analyst, we use the latest one in the restricted forecasting period.

Excess Market Return: Monthly return in excess of the value-weighted CRSP return and compounded over months in a quarter to estimate quarterly return.

DGTW Return: Monthly return minus the average return on stocks in DGTW segment style k to which stock i belongs and compounded over months in a quarter to estimate quarterly return.

4-Factor Alpha: Daily alpha estimated from time-series regressions of daily stock returns on Fama and French (1992) market, SMB and HML factors and Carhart (1997) UMD factor each month and compounded over days in a quarter to estimate quarterly return.

Cumulative Abnormal Return (CAR): CARs are computed from daily returns in excess of returns on a benchmark portfolio to which the stock belongs, over a three-day window around the earnings announcement dates [-1, 1]. Earnings announcement dates are obtained from I/B/E/S database. To construct benchmark daily returns, we follow French's website, and sort stocks into 2×3 benchmark portfolios by size (ME) and book-to-market equity (BE/ME) at June end of each year t. Median ME on NYSE stocks and the 30th and 70th percentiles of BE/ME on NYSE stocks, computed as book equity in the last fiscal year end in $t-1$ divided by ME in December of $t-1$, are used as breakpoints.

Table AII: Fund Size and Performance by Style Segments

Table reports fund size and performance by style segments. Portfolio formation is described in Table I. For each style segment, Panel A reports the average inflation adjusted total net assets under management (TNA) expressed in millions of dollars and in parentheses average 4-factor adjusted quarterly compounded gross alpha in percent. Panel B reports average gross value-added expressed in millions of dollars. Panel C reports regressions of gross alpha against ln TNA, regressions of gross value added against ln TNA and squared ln TNA, the characteristics of excessively underfunded, excessively over-funded and moderately funded funds. Gross alpha is estimated from rolling 12-month time series regressions of monthly fund gross return on Fama and French (1992) market, SMB, HML, and Carhart(1997) UMD factors, and as in Berk and Binsbergen (2015), averaged across current and prior months. Monthly gross fund return is the sum of fund monthly net return and one-twelfth of fund annual expense ratio. TNA is monthly total net assets under management averaged across current and prior months, adjusted by inflation. Gross value-added is the product of fund monthly gross alpha and TNA averaged across current and prior months.

		PANEL A Total Net Assets Under Management (4-Factor Quarterly Return Alpha)						PANEL B Gross Value-Added					
SIZE	BTM	MOMENTUM						MOMENTUM					
		1	2	3	4	5	***Size_BTM***	1	2	3	4	5	***Size_BTM***
1	1	594 (0.65)	543 (0.74)	532 (0.70)	520 (0.80)	581 (0.76)	554 (0.73)	4.05	4.09	4.08	3.05	3.33	3.72
1	2	598 (0.77)	538 (0.74)	501 (0.78)	483 (0.81)	487 (0.71)	521 (0.76)	3.05	3.22	2.52	3.10	2.86	2.95
1	3	628 (0.71)	535 (0.68)	470 (0.84)	482 (0.73)	457 (0.77)	514 (0.75)	3.97	4.02	3.17	2.34	2.62	3.23
1	4	573 (0.81)	460 (0.80)	457 (0.66)	468 (0.91)	453 (0.71)	482 (0.78)	3.98	2.73	2.41	3.52	2.39	3.01
1	5	544 (0.91)	520 (0.79)	497 (0.88)	469 (0.83)	468 (0.84)	500 (0.85)	2.91	2.64	3.20	3.07	2.78	2.92
2	1	596 (0.72)	561 (0.66)	548 (0.72)	575 (0.65)	617 (0.65)	580 (0.68)	4.10	3.34	3.64	3.71	3.89	3.74
2	2	595 (0.63)	585 (0.77)	534 (0.68)	574 (0.70)	590 (0.62)	575 (0.68)	3.69	3.98	3.10	3.54	3.10	3.48
2	3	574 (0.73)	542 (0.68)	509 (0.73)	552 (0.70)	544 (0.66)	544 (0.70)	3.42	3.95	3.36	3.14	2.98	3.37
2	4	563 (0.78)	511 (0.73)	530 (0.73)	489 (0.67)	507 (0.72)	520 (0.73)	3.62	3.09	3.60	3.11	3.30	3.34
2	5	567 (0.65)	576 (0.69)	541 (0.67)	541 (0.67)	508 (0.62)	546 (0.66)	3.67	4.07	3.51	2.73	3.15	3.42
3	1	629 (0.71)	626 (0.75)	587 (0.75)	614 (0.75)	643 (0.66)	620 (0.72)	3.92	3.65	3.93	3.93	4.02	3.89
3	2	641 (0.64)	611 (0.71)	573 (0.67)	572 (0.64)	615 (0.64)	602 (0.66)	4.21	0.56	2.57	3.35	1.20	2.38
3	3	629 (0.69)	580 (0.74)	566 (0.70)	575 (0.65)	612 (0.64)	592 (0.68)	4.53	3.31	3.25	3.32	0.72	3.03
3	4	631 (0.67)	595 (0.71)	583 (0.66)	577 (0.75)	585 (0.66)	594 (0.69)	3.63	0.17	3.45	3.66	0.39	2.26
3	5	637 (0.71)	631 (0.72)	630 (0.68)	583 (0.70)	646 (0.63)	625 (0.69)	3.80	3.50	3.83	1.66	3.78	3.32
4	1	837 (0.68)	802 (0.74)	780 (0.71)	723 (0.66)	710 (0.62)	770 (0.68)	5.82	5.43	4.36	4.77	3.87	4.85
4	2	849 (0.76)	764 (0.70)	766 (0.72)	749 (0.66)	726 (0.60)	771 (0.69)	1.22	-1.58	4.80	1.59	4.08	2.02
4	3	831 (0.68)	823 (0.69)	753 (0.70)	698 (0.67)	713 (0.70)	763 (0.69)	2.11	5.16	0.67	4.39	0.58	2.58
4	4	829 (0.70)	773 (0.64)	751 (0.68)	748 (0.73)	722 (0.66)	765 (0.68)	3.71	4.67	1.64	4.32	4.47	3.76
4	5	848 (0.70)	792 (0.68)	758 (0.76)	782 (0.69)	746 (0.65)	785 (0.70)	5.18	4.53	4.87	4.81	4.21	4.72
5	1	843 (0.60)	813 (0.62)	804 (0.61)	817 (0.60)	809 (0.58)	817 (0.60)	-0.37	-0.20	-0.50	-0.68	-0.21	-0.39

Table AII (Cont.)

		PANEL A Total Net Assets Under Management (4-Factor Quarterly Return Alpha)						PANEL B Gross Value-Added					
SIZE	BTM	MOMENTUM						MOMENTUM					
		1	2	3	4	5	***Size_BTM***	1	2	3	4	5	***Size_BTM***
5	2	817 (0.58)	812 (0.57)	797 (0.57)	797 (0.56)	803 (0.57)	805 (0.57)	-0.26	-0.57	-0.48	-0.13	-0.49	-0.39
5	3	827 (0.58)	807 (0.60)	800 (0.56)	800 (0.6)	820 (0.59)	811 (0.58)	-0.19	-0.15	-0.57	-0.07	-0.59	-0.31
5	4	843 (0.62)	827 (0.61)	800 (0.60)	808 (0.54)	820 (0.57)	820 (0.59)	-0.09	0.47	-0.36	-0.15	-0.88	-0.20
5	5	869 (0.59)	850 (0.59)	857 (0.55)	832 (0.58)	824 (0.57)	846 (0.58)	0.75	-0.94	-0.86	0.48	-2.08	-0.53

PANEL C

	Gross Alpha	Gross Value Added	Zhu (2018)
$ln(TNA)$	-0.0049^{c}	14.1149^{c}	-0.002^{c}
	(0.000)	(0.000)	(0.000)
$[ln(TNA)]^2$	-	-2.4575^{c}	-
		(0.000)	
Constant	0.0391^{c}	-	0.003
	(0.000)		
NOBS	1,800	1,800	
Adjusted R^2	0.907	0.623	
Quarter Fixed Effects	Y	Y	

			Gross Value Added ($mil)			Gross Alpha (bps)			TNA		
	Zhu (2018)	Percentile	Average (σ)	Min	Max	Average (σ)	Min	Max	Average (σ)	Min	Max
Overfunded	58.8	31.20	-45.0	-90.0	0.0	-0.0045	-0.0070	-0.0021	685.5	417.5	953.6
			(63.6)			(0.0035)			(379.1)		
Moderately Funded	25.9	43.60	3.6	0.0	7.1	0.0040	-0.0040	0.0120	619.7	531.5	707.9
			(5.0)			(0.0113)			(233.1)		
Underfunded	17.5	25.20	28.6	7.1	50.2	0.0260	0.0089	0.0430	1,218.0	861.1	1,574.9
			(30.4)			(0.0241)			(504.7)		

Table AIII: Forecast Returns: Fund Performance using Industry Concentration

Table reports style and quarter fixed effects regressions of lead quarter buy-and-hold stock returns on investment quality and control variables. ‡indicates that industry concentration are used to proxy fund performance. The cross-product of active holding and industry concentration summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. Average quarterly returns are expressed in percent. Variable definitions can be found in Table II. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. In two-way fixed effects regressions, errors are clustered by style and quarter. p-values are reported in parentheses. Superscript $^{a, b, c}$ denote statistical significance at 10%, 5%, and 1% levels.

	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ^{\ddagger} \equiv \theta_i^s$	1.255^{c}	1.295^{c}	1.131^{c}	1.097^{c}	1.337^{c}	1.307^{c}	1.146^{c}	1.100^{c}	0.881^{c}	1.036^{c}	0.949^{c}	0.879^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ^2 \equiv \theta_i^{s^2}$	-1.087^{c}	-1.161^{c}	-0.933^{c}	-0.904^{c}	-1.176^{c}	-1.188^{c}	-0.977^{c}	-0.920^{c}	-0.737^{c}	-0.913^{c}	-0.765^{c}	-0.700^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\Delta Breadth$	0.345^{b}	0.166	0.032	-0.033	0.363^{c}	0.182^{a}	0.076	0.005	0.261^{b}	0.204^{b}	0.050	-0.006
	(0.024)	(0.115)	(0.694)	(0.627)	(0.005)	(0.051)	(0.275)	(0.940)	(0.021)	(0.015)	(0.419)	(0.911)
$\Delta Active\ MF\ Ownership$	-0.037	-0.081	-0.034	-0.008	-0.058	-0.067	-0.041	-0.020	-0.045	-0.086	-0.019	0.003
	(0.543)	(0.161)	(0.544)	(0.873)	(0.357)	(0.154)	(0.394)	(0.642)	(0.498)	(0.108)	(0.709)	(0.953)
Holding Dispersion Index	0.037	-0.035	-0.031	-0.005	-0.008	-0.040	-0.034	-0.007	0.045	-0.047	-0.032	0.007
	(0.426)	(0.343)	(0.367)	(0.860)	(0.865)	(0.265)	(0.275)	(0.773)	(0.298)	(0.175)	(0.279)	(0.770)
Ln(MCAP)	-0.736	-0.464	-0.263	-0.166	-0.804	-0.375	-0.135	0.034	-0.619	-0.468	-0.471	-0.414
	(0.193)	(0.252)	(0.420)	(0.537)	(0.120)	(0.293)	(0.626)	(0.874)	(0.356)	(0.380)	(0.261)	(0.262)
Ln(Book-to-Market)	0.350	0.465	0.562^{b}	0.551^{c}	0.010	0.187	0.275	0.289^{b}	-0.001	0.209	0.332^{b}	0.375^{c}
	(0.419)	(0.129)	(0.019)	(0.007)	(0.973)	(0.381)	(0.104)	(0.041)	(0.996)	(0.220)	(0.028)	(0.004)
Prior Year Return	-0.108	-0.120	-0.191	-0.204	-0.094	-0.055	-0.094	-0.088	0.100	0.047	0.003	-0.010
	(0.775)	(0.680)	(0.436)	(0.312)	(0.751)	(0.772)	(0.549)	(0.503)	(0.490)	(0.680)	(0.974)	(0.909)
CRSP Turnover	-0.306	-0.538^{c}	-0.527^{c}	-0.535^{c}	-0.120	-0.292^{b}	-0.300^{c}	-0.337^{c}	-0.305	-0.525^{c}	-0.571^{c}	-0.627^{c}
	(0.106)	(0.000)	(0.000)	(0.000)	(0.434)	(0.027)	(0.007)	(0.001)	(0.200)	(0.006)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.071	-0.178	-0.270	-0.272	-0.489	-0.622^{b}	-0.661^{c}	-0.599^{c}	-0.449	-0.704^{b}	-0.797^{c}	-0.771^{c}
	(0.905)	(0.692)	(0.494)	(0.389)	(0.189)	(0.020)	(0.004)	(0.002)	(0.158)	(0.011)	(0.001)	(0.000)
Market Beta	-0.175	-0.190	-0.154	-0.172	-0.067	-0.107	-0.085	-0.096	-0.341^{c}	-0.396^{c}	-0.296^{c}	-0.270^{c}
	(0.560)	(0.378)	(0.329)	(0.221)	(0.719)	(0.477)	(0.429)	(0.322)	(0.004)	(0.000)	(0.001)	(0.001)
NOBS	182,023	179,098	176,206	173,359	181,908	178,875	175,860	172,903	184,524	181,728	178,883	176,080
R^2	0.036	0.042	0.044	0.046	0.005	0.012	0.019	0.024	0.008	0.016	0.022	0.028

Table AIV: Regression Results on Control Variables

Table IV (contd.)												
Ln(MCAP)	-0.578[c]	-0.275	-0.128	-0.031	-0.264[c]	0.096	0.218[b]	0.292[c]	-0.503[b]	-0.259	-0.137	-0.067
	(0.006)	(0.120)	(0.409)	(0.826)	(0.009)	(0.253)	(0.011)	(0.001)	(0.028)	(0.180)	(0.408)	(0.677)
Ln(Book-to-Market)	0.036	0.078	0.125	0.150	-0.125	-0.035	0.030	0.060	0.068	0.115	0.160	0.177
	(0.877)	(0.725)	(0.544)	(0.404)	(0.313)	(0.782)	(0.798)	(0.553)	(0.606)	(0.408)	(0.267)	(0.186)
Prior Year Return	0.038	0.048	0.019	0.026	-0.030	0.007	0.008	0.038	0.092	0.072	0.053	0.080
	(0.890)	(0.831)	(0.923)	(0.880)	(0.852)	(0.946)	(0.925)	(0.631)	(0.377)	(0.413)	(0.573)	(0.419)
CRSP Turnover	-0.162	-0.340[b]	-0.390[c]	-0.430[c]	-0.175	-0.327[b]	-0.348[c]	-0.380[c]	-0.250	-0.422[c]	-0.472[c]	-0.506[c]
	(0.410)	(0.042)	(0.008)	(0.001)	(0.260)	(0.014)	(0.003)	(0.000)	(0.104)	(0.002)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.725[b]	-0.989[c]	-1.020[c]	-0.955[c]	-0.635[c]	-0.844[c]	-0.868[c]	-0.799[c]	-0.734[c]	-1.055[c]	-1.132[c]	-1.161[c]
	(0.041)	(0.001)	(0.000)	(0.000)	(0.007)	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)
Market Beta	0.097	0.033	-0.011	-0.080	0.119	0.067	0.038	-0.014	-0.499[c]	-0.551[c]	-0.463[c]	-0.449[c]
	(0.625)	(0.846)	(0.937)	(0.567)	(0.475)	(0.638)	(0.746)	(0.905)	(0.005)	(0.000)	(0.000)	(0.000)
Constant	0.192	-0.695	-1.055[b]	-1.218[c]	-0.256	-1.155[c]	-1.443[c]	-1.560[c]	0.488[c]	-0.448[b]	-0.788[c]	-0.993[c]
	(0.699)	(0.133)	(0.010)	(0.001)	(0.438)	(0.000)	(0.000)	(0.000)	(0.001)	(0.010)	(0.000)	(0.000)

Table IX Panel A (contd.)												
Market Capitalization	-0.827	-0.395	-0.290	-0.209	-1.001[a]	-0.459	-0.219	-0.061	-0.668	-0.437	-0.442	-0.398
	(0.131)	(0.333)	(0.372)	(0.444)	(0.052)	(0.195)	(0.425)	(0.778)	(0.324)	(0.435)	(0.326)	(0.325)
Book-to-Market	0.674	0.824[b]	0.792[c]	0.791[c]	0.110	0.283	0.329[a]	0.370[b]	-0.024	0.231	0.357[b]	0.432[c]
	(0.176)	(0.015)	(0.002)	(0.001)	(0.739)	(0.201)	(0.059)	(0.014)	(0.914)	(0.178)	(0.021)	(0.002)
Prior Year Return	-0.585	-0.536	-0.507	-0.497[a]	-0.307	-0.242	-0.238	-0.233	0.119	-0.009	-0.063	-0.112
	(0.221)	(0.169)	(0.100)	(0.060)	(0.316)	(0.241)	(0.158)	(0.104)	(0.464)	(0.938)	(0.528)	(0.222)
CRSP Turnover	-0.215	-0.442[c]	-0.441[c]	-0.465[c]	-0.083	-0.263[b]	-0.265[b]	-0.309[c]	-0.271	-0.441[b]	-0.494[c]	-0.544[c]
	(0.257)	(0.004)	(0.000)	(0.000)	(0.576)	(0.036)	(0.014)	(0.002)	(0.251)	(0.019)	(0.001)	(0.000)
Idiosyncratic Volatility	-0.432	-0.617	-0.621	-0.567[a]	-0.663[a]	-0.812[c]	-0.828[c]	-0.743[c]	-0.524	-0.862[c]	-0.968[c]	-0.944[c]
	(0.493)	(0.198)	(0.130)	(0.093)	(0.088)	(0.002)	(0.000)	(0.000)	(0.104)	(0.002)	(0.000)	(0.000)
Market Beta	-0.386	-0.328	-0.275[a]	-0.291[a]	-0.164	-0.148	-0.131	-0.147	-0.348[c]	-0.398[c]	-0.301[c]	-0.287[c]
	(0.260)	(0.159)	(0.094)	(0.059)	(0.433)	(0.336)	(0.207)	(0.134)	(0.003)	(0.000)	(0.001)	(0.000)

Table X Panel A (contd.)

Market Capitalization	-1.244[b]	-0.878[b]	-0.632[a]	-0.506[a]	-1.268[b]	-0.728[b]	-0.447	-0.263	-0.916	-0.705	-0.698[a]	-0.619[a]
	(0.033)	(0.036)	(0.059)	(0.066)	(0.013)	(0.042)	(0.110)	(0.229)	(0.169)	(0.181)	(0.087)	(0.081)
Book-to-Market	0.372	0.480	0.574[b]	0.559[c]	0.028	0.197	0.284[a]	0.293[b]	0.012	0.215	0.337[b]	0.378[c]
	(0.390)	(0.119)	(0.017)	(0.007)	(0.924)	(0.356)	(0.096)	(0.040)	(0.955)	(0.206)	(0.026)	(0.004)
Prior Year Return	-0.081	-0.096	-0.167	-0.181	-0.068	-0.034	-0.074	-0.067	0.116	0.062	0.020	0.006
	(0.831)	(0.740)	(0.492)	(0.369)	(0.819)	(0.859)	(0.639)	(0.612)	(0.422)	(0.582)	(0.840)	(0.946)
CRSP Turnover	-0.347[a]	-0.567[c]	-0.552[c]	-0.555[c]	-0.151	-0.313[b]	-0.318[c]	-0.351[c]	-0.328	-0.536[c]	-0.582[c]	-0.637[c]
	(0.068)	(0.000)	(0.000)	(0.000)	(0.317)	(0.016)	(0.004)	(0.001)	(0.169)	(0.005)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.092	-0.200	-0.292	-0.289	-0.518	-0.648[b]	-0.685[c]	-0.618[c]	-0.463	-0.725[c]	-0.816[c]	-0.788[c]
	(0.877)	(0.658)	(0.464)	(0.362)	(0.164)	(0.016)	(0.003)	(0.001)	(0.147)	(0.009)	(0.001)	(0.000)
Market Beta	-0.187	-0.199	-0.160	-0.177	-0.075	-0.113	-0.087	-0.099	-0.343[c]	-0.396[c]	-0.295[c]	-0.270[c]
	(0.528)	(0.351)	(0.308)	(0.205)	(0.683)	(0.449)	(0.415)	(0.307)	(0.004)	(0.000)	(0.001)	(0.000)

Table X Panel B (contd.)

Market Capitalization	-1.160[b]	-0.802[a]	-0.570[a]	-0.451	-1.186[b]	-0.653[a]	-0.384	-0.210	-0.849	-0.644	-0.644	-0.573
	(0.045)	(0.055)	(0.088)	(0.101)	(0.020)	(0.069)	(0.171)	(0.342)	(0.202)	(0.223)	(0.116)	(0.110)
Book-to-Market	0.338	0.450	0.547[b]	0.535[c]	-0.005	0.169	0.259	0.272[a]	-0.011	0.193	0.316[b]	0.360[c]
	(0.435)	(0.143)	(0.023)	(0.009)	(0.987)	(0.431)	(0.130)	(0.057)	(0.958)	(0.256)	(0.035)	(0.005)
Prior Year Return	-0.083	-0.098	-0.169	-0.182	-0.070	-0.036	-0.075	-0.068	0.113	0.060	0.018	0.005
	(0.826)	(0.734)	(0.490)	(0.367)	(0.812)	(0.850)	(0.633)	(0.608)	(0.433)	(0.595)	(0.853)	(0.957)
CRSP Turnover	-0.305	-0.528[c]	-0.518[c]	-0.524[c]	-0.111	-0.277[b]	-0.286[b]	-0.323[c]	-0.300	-0.508[c]	-0.556[c]	-0.613[c]
	(0.107)	(0.000)	(0.000)	(0.000)	(0.465)	(0.035)	(0.010)	(0.002)	(0.208)	(0.007)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.080	-0.188	-0.280	-0.278	-0.507	-0.638[b]	-0.675[c]	-0.608[c]	-0.457	-0.718[c]	-0.809[c]	-0.781[c]
	(0.893)	(0.678)	(0.482)	(0.382)	(0.175)	(0.018)	(0.004)	(0.001)	(0.152)	(0.010)	(0.001)	(0.000)
Market Beta	-0.182	-0.196	-0.157	-0.175	-0.071	-0.109	-0.084	-0.096	-0.339[c]	-0.393[c]	-0.292[c]	-0.268[c]
	(0.540)	(0.361)	(0.317)	(0.212)	(0.702)	(0.466)	(0.431)	(0.320)	(0.004)	(0.000)	(0.001)	(0.001)

Table XI (contd.)												
Market Capitalization	-1.232[b]	-0.853[b]	-0.602[a]	-0.478[a]	-1.251[b]	-0.707[b]	-0.421	-0.239	-0.896	-0.678	-0.667	-0.589[a]
	(0.034)	(0.042)	(0.073)	(0.083)	(0.014)	(0.049)	(0.133)	(0.276)	(0.181)	(0.201)	(0.104)	(0.099)
Book-to-Market	0.372	0.480	0.574[b]	0.558[c]	0.028	0.197	0.283[a]	0.293[b]	0.012	0.215	0.336[b]	0.378[c]
	(0.390)	(0.119)	(0.017)	(0.007)	(0.924)	(0.357)	(0.097)	(0.041)	(0.955)	(0.207)	(0.026)	(0.004)
Prior Year Return	-0.079	-0.095	-0.166	-0.180	-0.066	-0.032	-0.073	-0.066	0.116	0.063	0.021	0.007
	(0.834)	(0.744)	(0.495)	(0.370)	(0.823)	(0.866)	(0.645)	(0.616)	(0.422)	(0.579)	(0.831)	(0.937)
CRSP Turnover	-0.348[a]	-0.568[c]	-0.554[c]	-0.557[c]	-0.154	-0.315[b]	-0.321[c]	-0.354[c]	-0.332	-0.540[c]	-0.587[c]	-0.641[c]
	(0.068)	(0.000)	(0.000)	(0.000)	(0.310)	(0.016)	(0.004)	(0.001)	(0.165)	(0.004)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.092	-0.199	-0.290	-0.287	-0.518	-0.648[b]	-0.684[c]	-0.617[c]	-0.462	-0.723[c]	-0.814[c]	-0.786[c]
	(0.877)	(0.659)	(0.466)	(0.365)	(0.164)	(0.016)	(0.003)	(0.001)	(0.148)	(0.009)	(0.001)	(0.000)
Market Beta	-0.188	-0.200	-0.160	-0.178	-0.076	-0.114	-0.088	-0.099	-0.344[c]	-0.396[c]	-0.295[c]	-0.271[c]
	(0.526)	(0.348)	(0.305)	(0.203)	(0.677)	(0.444)	(0.410)	(0.303)	(0.004)	(0.000)	(0.001)	(0.000)

Table XII (contd.)												
$\Delta Breadth$	0.368[b]	0.187[a]	0.049	-0.018	0.384[c]	0.200[b]	0.089	0.017	0.276[b]	0.220[c]	0.061	0.003
	(0.017)	(0.078)	(0.547)	(0.792)	(0.004)	(0.033)	(0.196)	(0.775)	(0.015)	(0.009)	(0.321)	(0.954)
$\Delta Active\ MF\ Ownership$	-0.035	-0.089	-0.043	-0.014	-0.057	-0.075	-0.048	-0.026	-0.041	-0.095[a]	-0.027	-0.003
	(0.580)	(0.129)	(0.450)	(0.767)	(0.397)	(0.126)	(0.323)	(0.560)	(0.562)	(0.083)	(0.612)	(0.948)
Holding Dispersion Index	0.008	-0.061[a]	-0.050	-0.021	-0.036	-0.061[a]	-0.048	-0.018	0.038	-0.052	-0.033	0.008
	(0.872)	(0.095)	(0.140)	(0.440)	(0.409)	(0.082)	(0.123)	(0.465)	(0.395)	(0.129)	(0.270)	(0.760)
Market Capitalization	-1.290[b]	-0.897[a]	-0.610	-0.455	-1.267[b]	-0.661[a]	-0.348	-0.138	-0.757	-0.508	-0.470	-0.395
	(0.044)	(0.058)	(0.104)	(0.133)	(0.022)	(0.087)	(0.245)	(0.556)	(0.264)	(0.342)	(0.245)	(0.244)
Book-to-Market	0.354	0.466	0.561[b]	0.544[c]	0.011	0.184	0.271	0.280[b]	-0.000	0.203	0.324[b]	0.366[c]
	(0.413)	(0.128)	(0.020)	(0.008)	(0.969)	(0.389)	(0.112)	(0.049)	(0.998)	(0.233)	(0.031)	(0.005)
Prior Year Return	-0.072	-0.093	-0.167	-0.182	-0.063	-0.037	-0.079	-0.075	0.104	0.046	0.000	-0.012
	(0.848)	(0.745)	(0.489)	(0.360)	(0.831)	(0.843)	(0.611)	(0.566)	(0.464)	(0.677)	(0.997)	(0.891)
CRSP Turnover	-0.331[a]	-0.550[c]	-0.532[c]	-0.535[c]	-0.133	-0.292[b]	-0.294[c]	-0.327[c]	-0.299	-0.501[c]	-0.546[c]	-0.602[c]
	(0.078)	(0.000)	(0.000)	(0.000)	(0.367)	(0.023)	(0.007)	(0.001)	(0.213)	(0.009)	(0.000)	(0.000)
Idiosyncratic Volatility	-0.078	-0.185	-0.284	-0.282	-0.508	-0.637[b]	-0.680[c]	-0.614[c]	-0.461	-0.731[c]	-0.817[c]	-0.792[c]
	(0.896)	(0.683)	(0.477)	(0.374)	(0.178)	(0.019)	(0.004)	(0.001)	(0.148)	(0.009)	(0.001)	(0.000)
Market Beta	-0.188	-0.199	-0.159	-0.176	-0.073	-0.108	-0.082	-0.092	-0.330[c]	-0.381[c]	-0.280[c]	-0.259[c]
	(0.525)	(0.350)	(0.308)	(0.208)	(0.688)	(0.464)	(0.440)	(0.337)	(0.006)	(0.000)	(0.001)	(0.001)

ONLINE APPENDIX

Table OI: Forecast Returns: Investment Quality of Stocks by Firm Absolute Forecasting Error

Table reports two-way fixed effects regressions of lead quarter buy-and-hold stock returns on the investment quality of large, median, small absolute forecast errors (AFE) stocks with positive earnings, as well on negative earning stocks. AFE is estimated as the absolute difference of mean analyst forecast for the period from realized earnings, scaled by realized earnings. Stocks are sorted into terciles each quarter by AFE. The cross-product of active holding and GVA summed across funds is used to identify selection skill. Investment quality $IQ \equiv \theta_i^s$ is the odds ratio of a stock's relative percentile rank, θ_i^s, on selection skill. Average quarterly returns are expressed in percent. Variable definitions can be found in Table III. All variables are normalized by their standard deviations across the sample period. Control variables are demeaned. Errors are clustered by style and quarter. p-values are reported in parentheses. Superscripts $^{a, b, c}$ denote statistical significance at 10%, 5%, and 1% levels. Unreported estimated coefficients on control variables are reported in an Appendix.

	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
$IQ \times Large\ AFE$	0.637^{c}	0.780^{c}	0.832^{c}	0.797^{c}	0.587^{c}	0.710^{c}	0.754^{c}	0.694^{c}	0.359^{b}	0.518^{c}	0.608^{c}	0.592^{c}
	(0.004)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.031)	(0.000)	(0.000)	(0.000)
$IQ \times Mid\ AFE$	1.567^{c}	1.442^{c}	1.260^{c}	1.095^{c}	1.456^{c}	1.328^{c}	1.152^{c}	0.959^{c}	1.255^{c}	1.134^{c}	1.007^{c}	0.884^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ \times Small\ AFE$	1.788^{c}	1.299^{c}	1.070^{c}	0.924^{c}	1.767^{c}	1.246^{c}	1.025^{c}	0.846^{c}	1.543^{c}	1.158^{c}	0.975^{c}	0.847^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$IQ \times Neg\ Earnings$	0.651^{c}	0.724^{c}	0.626^{c}	0.534^{c}	0.616^{c}	0.639^{c}	0.543^{c}	0.425^{c}	0.381^{a}	0.453^{c}	0.415^{c}	0.392^{c}
	(0.009)	(0.000)	(0.000)	(0.000)	(0.008)	(0.000)	(0.000)	(0.003)	(0.081)	(0.004)	(0.007)	(0.005)
$IQ^2 \equiv \theta_i^{s\,2}$	-0.739^{c}	-0.722^{c}	-0.615^{c}	-0.548^{c}	-0.691^{c}	-0.631^{c}	-0.537^{c}	-0.451^{c}	-0.574^{c}	-0.554^{c}	-0.497^{c}	-0.449^{c}
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
$\Delta Breadth$	0.398^{b}	0.200^{a}	0.048	-0.036	0.382^{c}	0.201^{b}	0.080	-0.005	0.296^{b}	0.227^{c}	0.074	-0.003
	(0.016)	(0.069)	(0.566)	(0.622)	(0.007)	(0.040)	(0.257)	(0.939)	(0.015)	(0.005)	(0.214)	(0.954)
$\Delta Active\ MF\ Ownership$	-0.062	-0.089	-0.030	0.014	-0.075	-0.069	-0.034	0.008	-0.069	-0.100^{a}	-0.026	0.005
	(0.373)	(0.130)	(0.555)	(0.754)	(0.308)	(0.162)	(0.444)	(0.838)	(0.382)	(0.086)	(0.567)	(0.901)
Holding Dispersion	0.024	-0.057	-0.052	-0.025	-0.028	-0.059	-0.055^{a}	-0.024	0.047	-0.047	-0.035	0.001
	(0.616)	(0.119)	(0.129)	(0.394)	(0.561)	(0.112)	(0.083)	(0.361)	(0.325)	(0.189)	(0.245)	(0.967)
Market Capitalization	-1.215^{a}	-0.874^{a}	-0.702^{a}	-0.537^{a}	-1.212^{b}	-0.716^{a}	-0.498^{a}	-0.253	-0.600	-0.307	-0.331	-0.259
	(0.055)	(0.060)	(0.057)	(0.072)	(0.025)	(0.055)	(0.082)	(0.263)	(0.296)	(0.493)	(0.326)	(0.385)
Book-to-Market	0.329	0.402	0.488^{b}	0.469^{b}	0.021	0.168	0.250	0.254^{a}	-0.048	0.065	0.165	0.202^{a}
	(0.429)	(0.172)	(0.030)	(0.013)	(0.944)	(0.430)	(0.129)	(0.060)	(0.789)	(0.647)	(0.184)	(0.060)
Prior Year Return	-0.041	-0.076	-0.110	-0.120	-0.040	-0.033	-0.046	-0.046	0.028	-0.020	-0.032	-0.025
	(0.912)	(0.785)	(0.650)	(0.551)	(0.886)	(0.857)	(0.776)	(0.731)	(0.852)	(0.859)	(0.752)	(0.776)
CRSP Turnover	-0.453^{b}	-0.628^{c}	-0.581^{c}	-0.584^{c}	-0.241	-0.361^{c}	-0.335^{c}	-0.360^{c}	-0.408^{a}	-0.568^{c}	-0.575^{c}	-0.624^{c}
	(0.027)	(0.000)	(0.000)	(0.000)	(0.105)	(0.004)	(0.002)	(0.000)	(0.063)	(0.001)	(0.000)	(0.000)
Idiosyncratic Volatility	0.197	-0.054	-0.210	-0.204	-0.187	-0.437	-0.550^{b}	-0.505^{c}	-0.146	-0.576^{b}	-0.766^{c}	-0.766^{c}

Table OI (Cont.)

	EXCESS MARKET RETURN				DGTW RETURN				4-FACTOR ALPHA			
	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr	Lead 1 Qtr	Lead 2 Qtr	Lead 3 Qtr	Lead 4 Qtr
	(0.754)	(0.906)	(0.608)	(0.528)	(0.632)	(0.110)	(0.023)	(0.010)	(0.625)	(0.017)	(0.000)	(0.000)
Market Beta	-0.201	-0.219	-0.179	-0.200	-0.086	-0.128	-0.097	-0.107	-0.251[a]	-0.342[c]	-0.265[c]	-0.243[c]
	(0.515)	(0.324)	(0.272)	(0.165)	(0.661)	(0.412)	(0.393)	(0.287)	(0.053)	(0.000)	(0.002)	(0.001)
NOBS	152598	150532	148350	146166	152485	150317	148017	145731	153599	152351	150288	148150
R^2	0.039	0.044	0.045	0.046	0.005	0.012	0.018	0.023	0.006	0.012	0.019	0.025

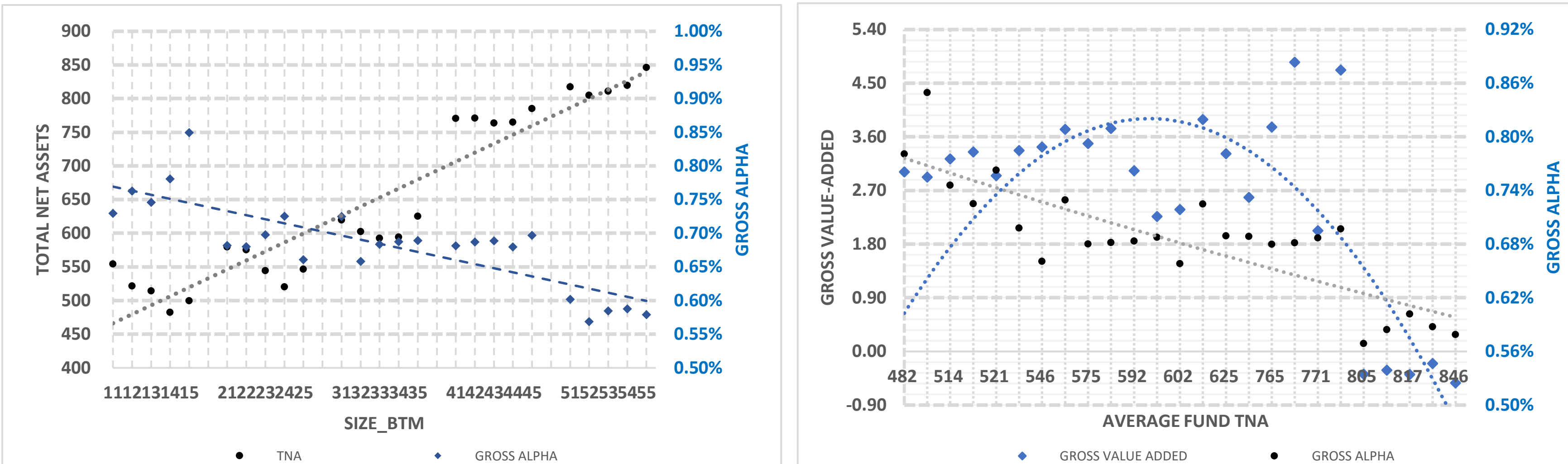


Figure I: Fund Size and Performance by Style Segment

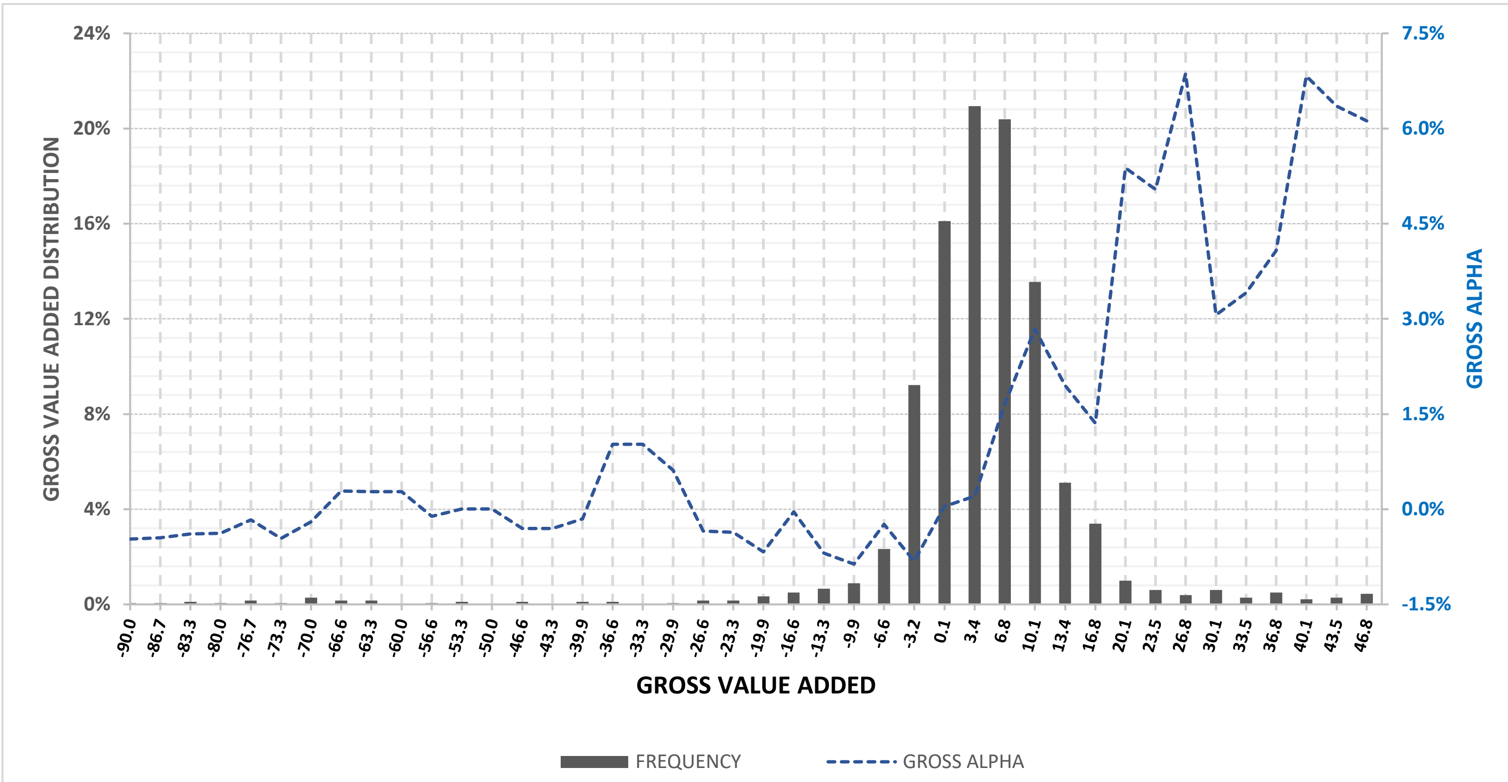


Figure II: Distribution of Gross Value-Added and Total Net Assets by Gross Alpha

Fund Flows and the Primary Market Activities of APs

Is There a Bright Side to ETFs

Cao Fang, Wayne Y. Lee[1], and Craig G. Rennie[2]

ABSTRACT

In this study, we use fund flow shocks from exogenous changes in ETF share demand to quantify the cost of trading stocks purchased or sold by APs in conjunction with the creation or redemption of ETF shares. We document a negative relation between return and the impact of primary market activities of APs on the liquidity of ETF-owned stocks. The stock-specific liquidity effect cannot be attributed to systematic asset pricing factors. Further, we find the improvements in liquidity from the primary market activities of APs enhance price discovery and strengthen the stock return-volatility relation.

Keywords: ETFs, Authorized Participants, Primary Market Activity, Liquidity, Price Discovery,

Asset Pricing

JEL classification: G10, G12, G14

---

[1] WLee@walton.uark.edu, 479-575-3944, Sam M. Walton College of Business, University of Arkansas, WCOB 475, Fayetteville, AR 72701, USA.

[2] crennie@walton.uark.edu, 501-819-2561, Sam M. Walton College of Business, University of Arkansas, WCOB 475, Fayetteville, AR 72701, USA.

I. Introduction

A number of empirical studies depict ETFs in a negative light. ETFs hold shares of stock in trust. The reduction both in the number of shares available for trade and in the volume of uninformed trading in secondary markets from a migration of retail investors to ETF shares lowers the incentive for informed trading on firm-specific information. Israeli, Lee, and Sridharan (2017) find an increase in bid-ask spreads and Amihud (2002) price impact with higher ETF ownership. Stock returns are also more correlated with market returns and future earnings impounded in current stock prices are more markedly discounted when ETF ownership is high. Further, Ben-David, Franzoni, and Moussawi (2018) find the intraday liquidity, convenience, and low transactions cost of ETF shares foster noise trading in ETF shares by retail investors and higher portfolio turnover by short-horizon institutional investors in stocks where ETF ownership is high. Changes in the price of stocks underlying ETFs from shocks to daily ETF fund flows are short-lived and deviations in ETF share price from NAV catalyze arbitrage trading by Authorized Participants (APs) and other high-frequency traders. Arbitrage activity transmits liquidity shocks in ETFs onto the underlying stocks. The resulting increase in intraday and daily stock volatility creates an undiversifiable risk.

Other studies present a more positive view of ETFs. Glosten, Nallareddy, and Zhu (2019) find that ETF activity improves price discovery. ETFs are baskets of securities that confer trading advantages for retail and institutional investors. The intraday trading of ETFs mitigates the adverse selection cost of trading in individual stocks by retail investors and facilitates factor informed trading by institutional investors. ETF activity is more likely to reflect market and industry information than firm-level idiosyncratic information. ETF activities do not predict future stock returns but improve short-run informational efficiency. Systematic aggregate earnings information is incorporated into current stock returns in a more timely manner on more opaque small capitalization stocks and stocks with low analyst following.

The impact of the primary market activities of APs in response to exogenous demand shocks in ETF shares on the liquidity of ETF-owned stocks is relatively unexplored in current literature. In this study, we address two questions that frame the sources of the divergent views on ETFs. Do the primary market activities of APs make liquidity on ETF-owned stocks better or worse? Do the primary market activities of APs strengthen or weaken the relationship between stock return and volatility?

The primary market activities of APs will improve liquidity when the transaction costs associated with the purchases or sales of stocks in conjunction with the creation and redemption of ETF shares are low. We derive a simple aggregate fund flow-driven transaction cost assuming the average cost of trading a dollar of stock is proportional to stock turnover, the dollar volume of shares traded as a percent of aggregate fund ownership, scaled by the market capitalization of the stock. We sum the cost of trading the individual stocks across APs on ETFs that own the same stock. The total cost of trading a stock will be lower and liquidity will be higher on stocks when aggregate ETF ownership may be high but small relative to the market capitalization of the stock, aggregate ETF ownership is diffuse across a large number of ETFs, and the volatility of ETF fund flow shocks are less correlated across ETFs who own the stock.

Expected returns will be lower when the primary activities of APs improve liquidity on ETF owned stocks. The decreases in expected returns from increases in the trading volume of APs on ETF-owned stocks is a stock-specific liquidity effect distinct from and unexplained by market, size, book-to-market, momentum, investment, or profitability asset pricing factors. Transparent stocks will benefit least from the primary market activities of APs. Moreover, to the extent the intraday liquidity of ETFs attract investors to own stocks they would otherwise write off, the primary market activities of APs will be more important in smaller ETFs that specialize in niche stocks. The primary market activities of APs do not convey private information about future returns. Rather that publicly disclosed firm-specific information will be impounded into current stock returns more efficiently. In this context, the primary market activities of APs enhance price discovery.

Stocks are embedded real options. Firms have the option to expand, contract, switch between, or abandon existing lines of business, as well as the option to wait or invest in new products or services. Equity value is a convex function of the underlying assets of a firm. Stock return and volatility will be positively related, and the unobservability of the process of returns on the firm's real options creates information uncertainty that deters participation by ambiguity averse investors. Improvements in the liquidity of ETF-owned shares from the primary market activities of APs will strengthen the stock return-volatility relation.

In our empirical analysis, we use monthly fund flow shocks from exogenous changes in the demand for ETF shares to compute the monthly cost of trading the underlying stocks across APs on ETFs that own the same stock. The impact of primary market activities on liquidity is high when the cost of trading the stock is low, and conversely, is low when the cost of trading the stock is high. We sort stocks into low and high quintiles of liquidity each month.

Using Fama and MacBeth (1973) cross-sectional regressions, we find improvements in liquidity from the primary market activities of APs decrease stock returns. Monthly stock returns in the lowest liquidity quintile are 0.857% higher and 1.068% lower in the highest liquidity quintile than the monthly average 0.987% return. When ETF ownership and changes in ETF ownership are also taken into account, the spread between the highest and lowest liquidity quintile decreases from 1.925% to 1.179% but remains statistically and economically significant..

As expected, returns are higher on illiquid stocks. A one percent increase in Amihud illiquidity will increase stock return by 2.23%, and by 2.35% when ETF ownership and change in ETF ownership are taken into account. As in Ben-David et al. (2018) and Glosten et al. (2019), we find a significant positive correlation between stock returns and ETF ownership and changes in ETF ownership of 0.031 and 0.019 respectively.

Using time-series regressions, we find the monthly differences in stock returns from average in the lowest and highest liquidity quintiles is an ETF stock-specific factor unrelated to and unexplained by monthly systematic risk factors described in the CAPM, Fama and French (1992) 3-factor, Carhart (1997) 4-factor, and Fama and French (2014) 5-factor models. The liquidity impact of the primary market activities of APs is an independent factor that explains cross-sectional differences in stock returns on ETF-owned stocks.

We find the impact of the primary activities of APs on liquidity is more important in opaque stocks. The difference in stock returns between the lowest and highest liquidity quintile is positive but insignificant on S&P500 stocks but a positive and significant 1.217% on non-S&P stocks. Similarly, we find the liquidity from the primary market activities of APs are more important for smaller ETFs that specialize in niche stocks. The difference in liquidity quintile spreads between the smallest and largest ETFs is a positive 0.224% though statistically insignificant.

We also find the primary market activities of APs on liquidity are more economically significant when unexpected earnings surprises are negative. When actual earnings miss forecast, the spread between the highest and lowest liquidity quintiles of 1.304% is almost twice the 0.670% spread when actual earnings beat forecast. Moreover, the primary market activities of APs on liquidity attenuate stock price declines associated with negative earnings surprises. The reduced likelihood of post earnings drifts from investor overreactions to earnings surprises supports the view that ETF activities enhance information efficiency.

As expected, there is a significant positive relation between stock return and volatility when stocks are embedded real options. A one percent change in asset volatility will change stock return by 0.188%, and by 0.09% from a percentage change in equity return volatility. Idiosyncratic risk is priced when information uncertainty deters ambiguity averse investors. The correlation between stock return and idiosyncratic return volatility of 0.094 is statistically and economically significant. As a proxy for the idiosyncratic volatility in the liquidity of ETF-owned stocks, the correlation between stock return and standard deviation of turnover of 0.094 is also statistically and economically significant.

The impact of the primary market activities of APs on liquidity strengthens the positive stock return-volatility relation. The stock return and volatility relation is unaffected in the lowest liquidity quintile. But in the highest liquidity quintile, the stock return and volatility relation is stronger. In the highest liquidity quintile, a one percent change in asset volatility will change stock return further by 1.125%, and by 0.619% from a one percent change in equity return volatility. The primary activities of APs do not increase volatility. Rather, that stock returns become more sensitive to volatility.

Lastly, the primary market activities of APs on liquidity also intensifies the adverse effect of information uncertainty on the participation of ambiguity averse investors. In the highest liquidity quintile, the correlation between stock return and idiosyncratic return volatility is a positive 0.039, and negative 0.012 in the lowest liquidity quintile. Similarly, in the highest liquidity quintile, the correlation between stock return and standard deviation of turnover is a positive 0.027, and negative 0.011 in the lowest liquidity quintile.

Our study contributes to two strands of literature. The primary market activities of APs establish a link between the liquidity of ETFs and their underlying stocks. In a theoretical model, Cespa and Foucault's (2014), show that liquidity improves when prices are more informative. The resulting improvement in cross-asset learning strengthens the incentive of market makers to provide liquidity across correlated assets. ETF share and underlying stock liquidity are positively correlated when ETF share and underlying stock prices are informative. Liquidity and price discovery are intertwined.

Arbitrage trading by APs do not affect the volatility in stock returns. Using minute-by-minute transactions, Box, Davis, Evans and Lynch (2021) do not find that intraday trading on ETFs have an effect on underlying stock returns. In Bae and Kim (2020), deviations of ETF share price from NAV are significant only when the underlying stocks are illiquid. In Iwadate (2021), deviations of NAV from

ETF share price are contagious when fund flow shocks across ETFs that own the same stock are correlated and aggregate common ownership is high. In Broman and Shum (2018), ETFs whose share prices are relatively more liquid than the underlying stocks attract positive net fund flows.

Second, the liquidity impact of the primary market activities of APs on the stock return-volatility relation complements the literature on institutional ownership and the commonality of liquidity. ETFs are baskets of securities. Exogenous fund flow shocks that lead to co-movements in underlying stock returns when common ownership is high increases the systematic risk in ETF-owned stocks. Systematic risk will be higher when improvements in liquidity from the primary activities of APs make the stock return-volatility relation stronger.

II. Fund Flow Measure of Liquidity

We use fund flow shocks from exogenous changes in the demand for ETF shares to quantify the volume of stocks purchased or sold by APs used in in-kind exchanges of ETF shares for stocks associated with the creation or redemption of ETF shares. APs are market makers in the underlying stocks as well as ETFs. Their actual costs of trading the underlying stocks are unobservable. We evaluate the liquidity impact of the primary market activities of APs from a projected cost of trading the underlying stocks assuming the average cost of trading a stock across APs will be higher when the aggregate volume of a stock traded by APs represents a larger percentage of the aggregate dollar value of the stock owned by APs scaled by the market capitalization of the stock.

A. Flow-driven Trading

For fund j, the value of its holdings of stock i in period t, is defined as the product of the number of shares held n_{ijt} and stock price p_{it}.

$$n_{ijt}p_{it} = w_{ijt}a_{jt} \tag{1}$$

where w_{ijt} denotes the percentage of the total net assets of fund, a_{jt}, allocated to stock i. From (1), the prior to current period change in the value of the holdings of stock i by fund j from additional purchases or sales in period t is:

$$\Delta n_{ijt}p_{it} = (n_{ijt} - n_{ijt-1})p_{it} \tag{2}$$

and fund flow, f_{jt}, is:[1]

$$f_{jt} = \sum_i (n_{ijt} - n_{ijt-1})p_{it} = a_{jt} - a_{jt-1} - a_{jt-1}R_{jt} \tag{3}$$

where $w_{ijt-1} = n_{ijt-1}p_{it-1}/a_{jt-1}$. From (1) and (2),[2] $\Delta n_{ijt}p_{it}$ describes the impact of fund flow shocks on stock trading by funds who own the stock.[3]

$$\begin{aligned}\Delta n_{ijt}p_{it} &= a_{jt}w_{ijt} - a_{jt-1}w_{ijt-1}(p_{it}/p_{it-1}) \\ &= w_{ijt-1}(a_{jt} - a_{jt-1}) + a_{jt}(w_{ijt} - w_{ijt-1}) - a_{jt-1}w_{ijt-1}r_{it} \\ &= \underbrace{w_{ijt-1}f_{jt}}_{liquidity} + \underbrace{a_{jt-1}w_{ijt-1}(R_{jt} - r_{ijt})}_{portfolio\ rebalancing} + \underbrace{[f_{jt} + a_{jt-1}(1 + R_{jt})]\Delta w_{ijt}}_{portfolio\ turnover} \\ &= \underbrace{w_{ijt-1}f_{jt}}_{passive} + \underbrace{\varepsilon_{ijt}}_{active} \qquad (4)\end{aligned}$$

where $R_{jt} = \sum_i w_{ijt-1}r_{ijt}$, and $r_{it-1} = \Delta p_{it}/p_{it-1}$. In (4), realizations of fund flow as well as stock and fund returns are stochastic. It is straightforward to show that $\sum_i \varepsilon_{ijt} = 0$.

The shocks in fund flow will embed changes in the liquidity needs of fund owners as well as systematic macroeconomic factors.[4] In (4), the changes in the ownership of stock i held by fund j can

---

[1] $$\begin{aligned}\boldsymbol{f_{jt}} &= \sum_i(\boldsymbol{n_{ijt} - n_{ijt-1}})\boldsymbol{p_{it}} = \sum_i \boldsymbol{n_{ijt}p_{it}} - \sum_i \boldsymbol{n_{ijt-1}p_{it-1}} - \sum_i \boldsymbol{n_{ijt-1}}\Delta\boldsymbol{p_{it}} = \boldsymbol{a_{jt} - a_{jt-1} - a_{jt-1}}\sum_i \boldsymbol{w_{ijt-1}r_{ijt}} \\ &= \boldsymbol{a_{jt} - a_{jt-1} - a_{jt-1}R_{jt}}.\end{aligned}$$

[2] In Greenwood and Thesmar (2011, p. 473), using a log-linearization of (2),

$$\begin{aligned}\Delta\boldsymbol{n_{ijt}p_{it}} &= \boldsymbol{w_{ijt}}(\boldsymbol{f_{jt}} + \sum_i \boldsymbol{n_{ijt}}\Delta\boldsymbol{p_{it}}) + \boldsymbol{a_{jt}w_{ijt}}\left\{\frac{\Delta\boldsymbol{w_{ijt}}}{\boldsymbol{w_{ijt}}} - \frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}}\right\} \\ &= \boldsymbol{w_{ijt}f_{jt}} + \boldsymbol{w_{ijt}a_{jt}}\sum_i \boldsymbol{w_{ijt}}\frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}} + \boldsymbol{a_{jt}}\Delta\boldsymbol{w_{ijt}} - \boldsymbol{a_{jt}w_{ijt}}\frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}} \\ &= \boldsymbol{w_{ijt}f_{jt}} + \boldsymbol{w_{ijt}}\sum_i \boldsymbol{n_{ijt}p_{it}}\frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}} + \boldsymbol{a_{jt}}\Delta\boldsymbol{w_{ijt}} - \boldsymbol{a_{jt}w_{ijt}}\frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}} \\ &= \boldsymbol{w_{ijt}f_{jt}} + \boldsymbol{w_{ijt}a_{jt}}\left[\sum_i\left(\boldsymbol{w_{ijt}}\frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}}\right) - \frac{\Delta\boldsymbol{p_{it}}}{\boldsymbol{p_{it}}}\right] + \boldsymbol{a_{jt}}\Delta\boldsymbol{w_{ijt}}\end{aligned}$$

and fund flow f_{jt} is defined as $a_{jt} - a_{jt-1} - \sum_i n_{ijt}\Delta p_{it}$ rather than as $a_{jt} - a_{jt-1} - \sum_i n_{ijt-1}\Delta p_{it}$.

[3] $$\begin{aligned}\Delta\boldsymbol{n_{ijt}p_{it}} &= \boldsymbol{a_{jt}w_{ijt} - a_{jt-1}w_{ijt-1}}(\boldsymbol{p_{it}/p_{it-1}}) = \boldsymbol{a_{jt}w_{ijt} - a_{jt-1}w_{ijt-1} - a_{jt-1}w_{ijt-1}}(\boldsymbol{p_{it}/p_{it-1}} - \boldsymbol{1}) \\ &= \boldsymbol{w_{ijt-1}}(\boldsymbol{a_{jt} - a_{jt-1}}) + \boldsymbol{a_{jt}}(\boldsymbol{w_{ijt} - w_{ijt-1}}) - \boldsymbol{a_{jt-1}w_{ijt-1}r_{it}} \\ &= \boldsymbol{w_{ijt-1}f_{jt}} + \boldsymbol{a_{jt}}\Delta\boldsymbol{w_{ijt}} + \boldsymbol{a_{jt-1}w_{ijt-1}R_{jt}} - \boldsymbol{a_{jt-1}w_{ijt-1}r_{it}} \\ &= \boldsymbol{w_{ijt-1}f_{jt}} + \boldsymbol{a_{jt-1}w_{ijt-1}}(\boldsymbol{R_{jt} - r_{ijt}}) + [\boldsymbol{f_{jt}} + \boldsymbol{a_{jt-1}}(\boldsymbol{1} + \boldsymbol{R_{jt}})]\Delta\boldsymbol{w_{ijt}}\end{aligned}$$

[4] In addition, investors also learn about the latent skill of active fund managers from past fund performance. Funds with superior recent performance will experience money inflows, while funds with poor performance will suffer outflows (Goldstein, Jiang, and Ng, 2017). A fund's past performance will also prompt idiosyncratic shocks to fund flow. The fund

be decomposed into a passive and active component. The passive component, the pro-rated dollar change in stock holdings from a shock in fund flow, $w_{ijt-1}f_{jt}$, is flow-driven liquidity trading. A one percent change in fund flow will result in a w_{ijt-1} per cent change in stock holdings.

The active component involves portfolio rebalancing, $a_{jt-1}w_{ijt-1}(R_{jt} - r_{ijt})$, and turnover $[f_{jt} + a_{jt-1}(1 + R_{jt})]\Delta w_{ijt}$. Portfolio rebalancing is the dollar change in stock holdings required to restore holdings to their original allocations $a_{jt-1}w_{ijt-1}$. The purchases of stocks whose returns underperform the overall fund return, $R_{jt} - r_{it} < 0$, and sales of stocks whose returns outperform the overall fund return, $R_{jt} - r_{it} > 0$, are independent of fund flow shocks.

Portfolio turnover is the dollar change in stock holdings from a change in allocation. The changes in allocation, Δw_{ijt}, are forward-looking. As Pastor, Stambaugh, and Taylor (2017) point out, the profit opportunities from private information motivate portfolio turnover. Portfolio turnover will be higher when stock mispricing is considerable, fund managers are more active and their portfolios are less diversified.[5] High portfolio turnover should predict higher future fund performance.

In short, passive flow-driven liquidity trading is uninformed and uncorrelated with the dollar changes in stock holdings from portfolio rebalancing and turnover, $E_t(w_{ijt-1}f_{jt} \cdot \varepsilon_{ijt}) = E_t(w_{ijt-1}f_{jt})E_t(\varepsilon_{ijt}) = 0$.[6] For passive ETFs that track an index, the change in holdings from rebalancing and turnover are zero.

B. Fund Flow-driven Liquidity

Trading cost is the conceptual basis of liquidity. If one trades the same dollar amounts of two

flow sensitivity to performance will greater the lower is the information cost to retail investors to learning about managerial skill (Huang, Wei and Yan, 2007), and learning about skill from past performance is less informative when realized portfolio returns are extreme (Franzoni and Schmalz, 2017).

[5]See Lee and Swaminathan (2000), Nagel (2005), and Pastor, Stambaugh, and Taylor (2022).

[6]But $E_t(w_{ijt}f_{jt+1} \cdot a_{jt-1}\Delta w_{ijt}) \neq 0$ for actively managed equity funds. The realized return from portfolio turnover in the current period can affect fund flow in the subsequent period. Managerial turnover following poor performance is intended to change the distribution of future fund returns. Lynch and Musto (2003) find for poor performing funds that change strategy, future fund flow and performance are less sensitive to current performance than poor performing funds that do not.

stocks, the stock with the lower trading cost has greater liquidity. Larger trades have higher proportional trading costs and price impact (Amihud, 2002; and Keim and Madhavan, 1997).

Given fund flow shock, f_{jt}, the passive dollar change in stock holdings is $w_{ijt-1} f_{jt}$. The total cost of fund flow-driven trading in stock i across funds is:

$$c_{it} = \sum_j \theta_i \left(w_{ijt-1} |f_{jt}| \right) \tag{5}$$

where θ_i is the average cost of a dollar of stock traded. We assume the average cost of a dollar of stock traded is proportional to stock turnover, $w_{ijt-1}|f_{jt}| / \sum_j w_{ijt-1} a_{jt-1}$, the dollar volume traded as a percent of aggregate fund ownership, scaled by the market capitalization of the stock, $mcap_{it}$.[7] The total cost of fund flow-driven trading in stock i across funds is:

$$c_{it} \propto \sum_j \left(\frac{1}{mcap_{it-1}} \cdot \frac{w_{ijt-1}|f_{jt}|}{\sum_j w_{ijt-1} a_{jt-1}} \right) \left(w_{ijt-1} |f_{jt}| \right) \tag{6}$$

We can modify and rewrite (6) as:

$$\begin{aligned} c_{it} &\propto \left(\frac{\sum_j w_{ijt-1} a_{jt-1}}{mcap_{it-1}} \right) \cdot \sum_j \left\{ \left(\frac{w_{ijt-1} a_{jt-1}}{\sum_j w_{ijt-1} a_{jt-1}} \right)^2 \left(f_{jt} / a_{jt-1} \right)^2 \right\} \\ &\propto own_{it-1} \cdot \sum_j \left[share_{ijt-1}^2 \left(f_{jt} / a_{jt-1} \right)^2 \right] \\ &\propto own_{it-1} \cdot \sum_j share_{ijt-1}^2 \cdot \sum_j \left[\omega_{ijt-1} \left(f_{jt} / a_{jt-1} \right)^2 \right] \\ &\propto own_{it-1} \cdot \sum_j share_{ijt-1}^2 \cdot E_* \left(f_{jt} / a_{jt-1} \right)^2 \end{aligned} \tag{7}$$

where $own_{it-1} = \sum_j w_{ijt-1} a_{jt-1} / mcap_{it-1}$ is aggregate fund ownership as a percent of the market capitalization of the stock, $mcap_{it-1}$, $share_{ijt-1} = w_{ijt-1} a_{jt-1} / \sum_j w_{ijt-1} a_{jt-1}$ is a fund's ownership as a percent of aggregate fund ownership, and E_* is an expectation taken with respect to a probability

[7] In Pastor, Stambaugh, and Taylor (2022), the cost of trading is applied to the fund's total net assets as if it were a stock. Fund liquidity is computed as:

$$\mathbf{c_{jt}} = \sum_{\mathbf{i}} \boldsymbol{\theta} \left(\frac{\mathbf{w_{ijt} a_{Jt}}}{\mathbf{mcap_{it}}} \right) \left(\mathbf{w_{ijt} a_{Jt}} \right) = \left(\boldsymbol{\theta} \mathbf{M_{jt}}^{-1} \mathbf{a_{jt}^2} \right) \left\{ \sum_{\mathbf{i}} \left[\mathbf{w_{ijt}}^2 / \left(\mathbf{mcap_{it}} / \mathbf{M_{jt}} \right) \right] \right\}$$

where total cost is scaled by $M_{jt} = \sum_i w_{ijt} mcap_{it}$, the total market capitalization of all stocks i owned by fund j, to make liquidity be comparable across funds,. Fund liquidity, liq_{jt}, is:

$$\boldsymbol{liq_{jt}} = \left\{ \sum_j \left[\boldsymbol{w_{ijt}}^2 / \left(\boldsymbol{mcap_{it}} / \boldsymbol{M_{jt}} \right) \right] \right\}^{-1}$$

measure $\omega_{ijt-1} = share_{ijt-1}^2 / \sum_j share_{ijt-1}^2$, a normed Herfindahl weight that reflects the concentration in stock ownership across funds.

From (7), we define the fund flow-driven cost of trading stock i, $\hat{c}_{it}$ as:

$$\hat{c}_{it} = \underbrace{own_{it-1}}_{\substack{fund\\ ownership}} \cdot \underbrace{\sum_j share_{ijt-1}^2}_{\substack{ownership\\ concentration}} \cdot \underbrace{\sqrt{Var_*(f_{jt}/a_{jt-1}) + E_*^2(f_{jt}/a_{jt-1})}}_{\substack{volatility\ of\\ fund\ flow}} \qquad (8)$$

In (8), the cost of trading will be low when ETF share ownership is a small percentage of the total number of shares outstanding of the stock, ETF ownership is diffuse, and the volatility of fund flow shocks are less correlated across ETFs who own the stock.[8] The liquidity impact of the primary activities of APs in response to fund flow shocks, $pmliq_{it}$, will be high when the cost of trading the stock, $\hat{c}_{it}$, is low, and conversely, will be low when the cost of trading is high. In subsequent analysis, we define $pmliq_{it}$ as the inverse of the cost of trading the stock, $\hat{c}_{it}^{-1}$.

C. PMLIQ and ETF Ownership

In the Ben-David et al. (2018) liquidity trading hypothesis, ETFs attract a new layer of liquidity demand. A rise in the frequency and magnitude of ETF demand shocks propagated by the primary activities of APs onto the underlying stocks increases stock return volatility. Using ETF ownership as a proxy for the trading impact of the primary market activities of APs, Ben-David et al. (2018) find a significant positive correlation between ETF ownership and stock return volatility of 0.077 for S&P500 stocks and 0.053 for Russell 3000 stocks. The correlation between ETF ownership and stock return is weaker on small stocks, they argue, because APs can minimize trading costs on ETFs that track indices with large capitalization stocks, but higher trading costs impede the propagation of fund

[8]Consider a linear projection of fund flow volatilities, $\left(f_{jt}/a_{jt-1}\right)^2 = \lambda_0 + \lambda_1\rho_{jt} + \lambda_3 R_{jt-1} + \varsigma_{jt}$, on correlations with macroeconomic recession and business condition factors ρ_{jt}, past fund performance R_{jt-1}, and an ortho
gonal sentiment (non-information) driven residual ς_{jt} with mean $E_*(\varsigma_{jt}) = 0$. Then $E_*\left(f_{jt}/a_{it-1}\right)^2 = \lambda_0 + \lambda_1 E_*(\rho_{jt}) + \lambda_3 E_*(R_{jt-1}) + E_*(\varsigma_{jt}^2)$. The impact of fund flow on stock illiquidity is least when the volatility in fund flow shocks are uncorrelated with either recession or business conditions, are less sensitive to either past fund performance or sentiment-driven trading, and thereby, are overall less correlated across funds.

flow shocks on ETFs that track small capitalization stocks. Further, higher stock returns from high ETF ownership suggest the increase in stock return volatility introduces an undiversifiable risk.

As (8) shows, ETF ownership captures only one of three factors that determine trading cost. The correlation between ETF ownership and trading cost can be negative when the other two factors are taken into account. When ETF ownership is high because a large number of ETFs own the stock, trading costs can be low when the distribution of ETF ownership is diffuse and fund flow shocks across ETFs are weakly correlated. The primary activities of APs can increase liquidity on high ETF ownership stocks.

Conversely, the primary market activities of APs can decrease liquidity on low ETF ownership stocks. When ETF ownership is low because a small number of ETFs own the stock, trading costs can be high when ETF ownership is concentrated in a few funds and fund flow shocks across the small number of ETFs are strongly correlated.

The propagation of fund flow shocks in ETF share demand by the primary activities of APs increases liquidity on the underlying stocks when trading costs are low.

<Insert Figures 1 and 2 here.>

As shown in Figure 1 and 2, the trading activities of APs have a liquidity and risk impact. Aggregate fund flow-driven cost of trading is negatively correlated with Amihud (2002) illiquidity and stock return volatility. The primary market activities of APs can lead to increases in liquidity and stock return volatility. Stock returns will be more sensitive to the volatility of the value of the firm's underlying assets when improvements in liquidity from the primary market activities of APs make prices of the underlying stocks adjust more promptly to new information.

III. Hypothesis

A. Baseline Hypothesis

First, when trading costs are low, the primary activities of APs in response to exogenous fund flow shocks in the demand for ETF shares channel liquidity onto the underlying stocks. The decreases in expected returns from increases in the trading volume of APs on ETF-owned stocks is a stock-specific liquidity factor distinct from and unexplained by systematic market, size, book-to-market, momentum, investment, or profitability asset pricing risk factors. Improvements in liquidity from the primary market activities of APs will impound new information into current prices more efficiently.

Second, we distinguish information asymmetry from information uncertainty. Investors are differentially informed when they agree on future states-of-the-world but disagree on the probability distribution of state-contingent payoffs. Information uncertainty arises when there is little consensus among investors about the possible states-of-the-world as well as its associated outcomes. Improvements in liquidity from the primary market activities of APs will heighten the stock return-volatility relation and intensify information uncertainty about the value of a firm's underlying assets that deter ambiguity averse investors.

In a related paper by Jiang, Lee, and Zhang (2005), differences in the quality of information about the probability distribution of future state-contingent payoffs result from imprecise private signals that investors receive. Framing, recency, and confirmation bias can lead investors to be overconfident, and positive feedback trading strategies can accentuate deviations from intrinsic value when public signals are also noisy.[9] Using a combination of firm age, stock return volatility, trading volume, and equity duration as proxies, Jiang et al. (2005) find that future returns are lower on stocks where investor overconfidence is high. Lower future returns that reflect reversals in price from overvaluation exhibit stronger momentum consistent with risky arbitrage when mispricing from informational cascades persist over long periods of time and arbitrageurs face a potential risk of ruin in the interim (Delong, Shleifer, Summers, and Waldmann, 1990).

[9]Jiang et al. (2005) use firm age, stock return volatility, and trading volume, singly and in combination, as their primary proxies for the information uncertainty associated with unobservable overconfidence.

Herein, we recognize that stocks embed real options on assets-in-place.[10] Equity value is a nonlinear (convex) function of the value of the underlying assets. Additionally, when there is scant evidence from history or experience to draw upon about the likelihood of success[11] and extrapolating future success from realized outcomes is challenging, the unobservability of the process of returns underlying a firm's assets creates information uncertainty.

When future states-of-the-world and the probability distribution of state-contingent outcomes are highly uncertain, investor ease or unease with best estimates will influence their decision whether or not to invest (Ellsberg, 1961). Information uncertainty will limit participation by ambiguity averse investors. As Merton (1987) shows, when investors choose to invest only in a subset of securities, there is a shadow price (opportunity cost) associated with the self-imposed constraint.[12] The shadow price of under-diversification will be positive and equal across investors who opt not to own the stock.[13] Idiosyncratic risk will be priced when the market for stocks is segmented. In equilibrium, expected returns will be higher when the primary market activities of APs accentuate asset volatility and information uncertainty deters ambiguity averse investors.[14]

Using a household survey, Dimmock, Kouwenberg, Mitchell, and Peijnenbug (2016) find, as theory predicts, that ambiguity aversion is negatively associated with stock market participation, the fraction of financial assets held in stocks, and foreign stock ownership, but positively related to own-company stock ownership. Conditional on stock ownership, ambiguity aversion is related to portfolio under-diversification, and during the financial crisis, ambiguity averse respondents were more likely to sell stocks.

---

[10]As McDonald and Siegel (1985) and Berk, Green, and Naik (1999) point out, the CAPM may explain expected returns on a firm's underlying assets but not necessarily the expected returns on its equity. Da, Guo, and Jagannathan (2012) show the presence of real options seems to explain the poor performance of the CAPM.

[11]Competence mitigates ambiguity aversion (Heath and Tversky, 1991).

[12]See Merton (1987), eq. 8b, p. 491.

[13]See Merton (1987), eq. 10, p. 491.

[14]See Merton (1987), eq. 16, p. 492.

B. Information Uncertainty Proxies

Using Black-Scholes (1973) and Merton (1974) option pricing models, we estimate the volatility of returns on real options embedded in the underlying assets of the firm following an iterative procedure outlined in Bharath and Shumway (2008). First, using historical data, we calculate a standard deviation of daily stock returns, σ_E, over an estimation window of T years, and take the book value of the firm's total liabilities at the end of the most recent quarter preceding the estimation window to be the face value of the firm's debt, D. Second, we infer the market value of the firm's underlying assets each day from the call option pricing formula, $E = AN(d_1) - Dexp(-r_f T)N(d_2)$, where E is the market value of equity, $d_1 = (\sigma_A \sqrt{T})^{-1}\{ln(A/D) + (r_f + \frac{1}{2}\sigma_A^2)T\}$ and $d_2 = d_1 - \sigma_A\sqrt{T}$. We initialize $\sigma_A = \sigma_E[E/(D+E)]$. Third, we calculate the log return on the market value of the firm's underlying assets each day over the estimation window to generate a new estimate of σ_A. The new estimate replaces the initial value to recompute another estimate of σ_A. The process is repeated until the absolute difference in adjacent values of σ_A is less than 10^{-3}.

Because equity value is a convex function of the value of the firm's underlying assets, stock returns and the volatility in stock returns will be positively correlated (Duffee, 1995). We also use the volatility and idiosyncratic volatility in stock returns as alternative proxies for information uncertainty and ambiguity aversion.

Ben-David et al. (2018) interpret a higher volatility in return on stocks with high ETF ownership as a consequence of an attraction of high frequency and liquidity traders to ETFs and the propagation of high frequency and liquidity trading onto underlying stocks by the primary activities of APs. Moreover, that the increase in volatility appears to introduce undiversifiable risk in prices because stocks with high ETF ownership earn a significant risk premium of up to 56 basis points monthly. The link between higher volatility and return through the primary activities of APs, however, can stem from the incidence of real options in stocks. We show in our subsequent analysis that in stocks where the primary market activities of APs are higher, the positive relation between stock volatility and return is stronger. The primary activities of APs do not increase stock volatility. Rather, that stock returns become more sensitive to volatility.

Lastly, there is an extensive literature that documents commonality in liquidity. In Chordia, Roll, and Subrahmanyam (2000), the commonality in liquidity stems from market-wide intertemporal trading responses to general price swings. Since trading volume is a principal determinant of dealer inventory, its variation will induce co-movements in optimal inventory levels which lead in turn to co-movements in individual bid-ask spreads, quoted depth, and other measures of liquidity. In other studies, commonality in liquidity arises from investor demand shocks on mutual fund ownership of stocks (Anton and Polk, 2014; Greenwood and Thesmar, 2011; Basak and Pavlova, 2013; and Lou, 2012). In Ben-David et al. (2018), arbitrage trading by APs transmit liquidity trading on ETF shares onto the underlying stocks. Da and Shive (2018) show that higher ETF arbitrage activity contributes to return co-movement at both the fund and the stock levels.

As Chordia, Subrahmanyam, and Anshuman (2001) point out, if liquidity affects stock returns, then stock returns should be positively correlated with turnover if investors care about the risk associated with fluctuations in liquidity.[15] We use the standard deviation in turnover as another proxy for idiosyncratic volatility in ETF-owned stocks.

IV. Data

A. ETF Sample

Using the CRSP mutual fund database, we construct a sample of domestic equity ETFs traded on major US exchanges over the period January 2002 through December 2019.[16] Following Ben-David et al. (2018), we identify ETFs using security type variables from the CRSP mutual fund database and

[15]See Chordia (2000).
[16]Sample starts in 2002 due to data availability.

share code of 73 from the CRSP database, and we restrict our sample to the following Lipper codes: CA, EI, G, GI, MC, MR, SG, SP, BM, CG, CS, FS, H, ID, NR RE, TK TL S, and UT. The selected Lipper codes cover both broad-based domestic equity ETFs and sector ETFs. We exclude leveraged ETFs and exchange traded notes (ETNs). Leveraged ETFs use futures and other derivatives to achieve leveraged exposure to U.S. equities, and ETNs involve fund sponsor risks that render them unsuitable for analysis in this study. To avoid survivorship bias, we allow the entry and exit of ETFs in our sample. Our final sample consists of 583 unique ETFs, ranging from 11 in January 2002 to 340 in December 2019.[17]

We obtain monthly returns, month-end assets under management, and fund flows on our ETF sample from CRSP and CRSP mutual fund databases. Monthly equity holdings on our ETF sample ETFs are from Morningstar Direct[SM], which includes voluntary (monthly) as well as required (quarterly) disclosures of portfolio holdings. We use ETF CUSIPs in Morningstar to match our sample ETFs to the CRSP mutual fund holdings database.

B. Stock Sample

Our sample contains 8,135 unique ETF-owned stocks from 2,835 in January 2002 to 3,471 in December 2019. We obtain daily and monthly closing share prices, volume, shares outstanding, returns, and other data on all common stocks from the CRSP database, but retain only those stocks traded on the NYSE, AMEX, or NASDAQ. Consistent with prior literature, we eliminate stocks with excessively low share prices (below $5) or equity market capitalization (less than $10 million at month end). Balance sheet data are obtained from the S&P Compustat database.

C. Descriptive Statistics

In Table I, we report summary statistics on the variables used in our analysis. The liquidity impact

[17]In Appendix Table II, we report the top 10 ETF advisory firms by number of funds, the distribution of average fund size (in millions of dollars) and number of unique stocks owned across ETFs over the entire sample period, and for the end (September 2017) and the beginning (June 2004) months.

of the primary market activities of APs, $PMLIQ$, is the inverse of the cost of trading defined by equation (8). $ETF\ Ownership$ is the total number of shares owned by ETFs at month end as a percentage of total shares outstanding. $\Delta ETF\ Ownership$ is the month-by-month change in ETF ownership for a given stock. Market capitalization is the market value of equity computed as closing price multiplied by total shares outstanding. Book-to-market is book equity to shareholders' equity calculated following Daniel and Titman (2006). Momentum is the cumulative monthly return in the preceding 12-month period, ending in 1 month prior to start of the month. Market beta is the estimated coefficient on a CAPM regression of excess stock returns, $r_{it} - r_{ft}$, on excess market returns, $r_{mt} - r_{ft}$, over rolling 36-month windows. Amihud illiquidity, $illiq_{i,d} = \sum_d |r_{i,d}| / dvol_{i,d}$, is computed as the absolute daily return divided by daily trading volume in millions of dollars averaged across days in the month. Turnover is French (2008) adjusted CRSP volume divided by shares outstanding. $Asset\ Volatility$ is estimated following Bharath and Shumway (2008) which captures the volatility of asset value for a given firm. $Stock\ Return\ Volatility$ is estimated as the standard deviation of daily returns of past 25 trading days. $Idio\_\ Stock\ Return\ Volatility$ is the standard deviation of residuals estimated from a CAPM regression of excess stock returns on excess market returns over rolling 36-month windows. $Stdev\_Turnover$ is the standard deviation of daily stock turnover over a month, with daily turnover estimated as daily trading volume scaled by shares outstanding.

< Insert Table II here. >

In Table II we report the end-of-month $PMLIQ$ and $ETF\ ownership$ for randomly drawn samples of firms across different industries from the beginning, middle, and end of our sample period. At the end of January 2002, July 2008, and January 2019 respectively in Panels A, B, and C. As expected, $PMLIQ$ is not strongly correlated with $ETF\ Ownership$. Moreover, $PMLIQ$ does not exhibit a significant industry pattern.

V. Primary Market Activity of Authorized Participants

We examine the effect of fund flow-driven trading on ETF shares on stock returns using Fama and MacBeth (1973) cross-sectional regressions. In the tables, we report the time-series averages of the estimated model parameters.

A. Primary Market Liquidity

Sorting stocks into quintiles each month by the impact of the primary market activities of APs on liquidity, $PMLIQ$, we estimate Fama-MacBeth (1973) cross-sectional regressions over our sample period 2000 to 2019. The dependent variable ret_i are monthly stock returns expressed in percent.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \sum_k \delta_k X_{ik} + \varepsilon_i \quad (9)$$

$PMLIQ_{i,q}$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard errors of the time-series average of estimated model parameters are computed using a 6-month lag.

The intercept α is the average monthly stock return, and coefficients β_q, the deviation in high and low quintile monthly stock returns from the average. Coefficients δ_k are the products of the standard deviation in stock return $\sigma(ret)$ and the correlations of control variables with stock return, $\rho(X_{ik}, ret_i)$. A one standard deviation change in a control variable will result in a $\delta_k/\sigma(ret)$ standard deviation change in stock return. Alternatively, a one percent change in a control variable will change stock return by $\delta_k/\sigma(X_k)$ percent. Sample estimates of $\sigma(ret)$ and $\sigma(X_k)$ are reported in Table I.

<Insert Table III here.>

Results are reported in Table III. As conjectured, the primary activities of APs increase liquidity and lower expected returns. Increases in liquidity from the primary market activities of APs do not predict future returns as evident in Columns 4 to 6. Taking asset pricing factors and stock illiquidity into account, Column 2 shows returns on stocks in the lowest quintile of primary market liquidity are 0.857% higher, and 1.068% lower in the highest quintile, than the average 0.987% return. In Column

3, when $ETF\ Ownership$ and $\Delta ETF\ Ownership$ are also taken into account, returns on stocks in the lowest quintile are 0.555% higher and 0.624% lower in the highest quintile than the average 0.95% return. Compared to average stock returns, the spreads between the highest and lowest quintile of primary market liquidity, of 1.925% and 1.179%, are statistically and economically significant.

As expected, returns are higher on illiquid stocks. In Column 2, one percent increase in Amihud illiquidity will increase stock return by 2.23% (=6.932/3.11), and in Column 3, by 2.35%. Returns are also higher on stocks with greater systematic market risk and on growth stocks. Lower returns on small capitalization and high momentum stocks suggest a considerable number of ETFs in our sample own less volatile large capitalization and low momentum stocks that have lower expected returns. Ben David et al. (2018, Table X Panel B) finds similar results.

In Column 3, a one standard deviation increase in $ETF\ Ownership$ will result in a 0.031 (=0.548/17.68) standard deviation increase in stock return. This increase in return from $ETF\ Ownership$ is considerably smaller than the 56 bps increase that Ben-David et al. (2018) attribute to an undiversifiable risk created by the propagation of high frequency and liquidity trading in ETF shares onto the underlying securities through arbitrage trading by APs. In Israeli et al. (2017), a higher stock return is a consequence of higher transaction costs when the lockup of shares in trust reduces the supply of shares available for trade. We resolve the conflicting views about the impact of $ETF\ Ownership$ on return in subsequent analysis.

Further, as Glosten et al. (2019) predicts, factor-informed investors will frequently trade both the ETF and the underlying securities but idiosyncratically informed investors are likely to trade only the security about which they have information.[18] ETF activity in a stock will reveal systematic market-

[18]As Cong and Zhu (2016) point out, factor speculators prefer composite securities because they can exploit their informational advantage without creating a large price impact in the primary asset market. Factor liquidity traders also prefer composite securities because collectively they face lower adverse selection costs of trading against informed asset speculators on a subset of primary assets. In this regard, composite securities are quintessentially a factor investing tool.

wide information. A one standard deviation increase in $\Delta ETF\ Ownership$ will lead to a 0.019 (=0.352/17.68) standard deviation increase in stock return.

To examine whether the liquidity impact of the primary market activities of APs on stock returns is independent of asset pricing factors, we estimate two time-series regressions over our sample period.

$$ret_t(PMLIQ\_q) = \alpha + \sum_k \delta_k X_{k,t} + \varepsilon_t \tag{10}$$

where $ret_t(PMLIQ\_q) \equiv \beta_{q,t}$ are the time-series of estimated average returns on stocks sorted into $q = \{PMLIQ_{Low}, PMLIQ_{High}\}$ quintiles in monthly cross-sectional regressions in (9), and $X_{k,t}$ are monthly risk factor returns in the CAPM, Fama and French (1992) 3-factor, Carhart (1997) 4-factor, and Fama and French (2014) 5-factor models. MKT-RF is a CRSP value-weighted return in excess of the risk free rate on stocks of US firms listed on the NYSE, AMEX, or NASDAQ exchanges. SMB and HML are value-weighted returns on long-short portfolios of small and large capitalization stocks and high and low book-to-market stocks. MOM is an equal-weighted return on a long-short portfolio of stocks with the highest and lowest prior 12-month cumulative return lagged one month. CMA and RWA are the average returns on long-short portfolios of stocks with the highest and lowest operating profitability and investment.

<Insert Table IV here.>

The time-series regression results are reported in Table IV. Controlling for asset pricing factors, the liquidity impact of the primary market activities of APs is an independent factor that explains cross-sectional differences in stock returns on ETF-owned stocks.

B. Transparency

The primary market activities of APs should be less important in transparent stocks. We use two proxies for transparency. First, membership in the S&P500 which consists of stocks that are large capitalization, widely held and followed. Second, membership in smaller ETFs that specialize in niche (opaque) stocks that investors would otherwise not own absent the intraday liquidity of the ETFs.

S&P500 Membership

Table V reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions described in (11) estimated each month over our sample period 2000 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + I_{S\&P} \cdot \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \sum_k \delta_k X_{ik} + \varepsilon_i \quad (11)$$

where $I_{S\&P}$ is equal to 1 if the stock is in the S&P500 and 0 otherwise. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard errors of the time-series average of estimated model parameters are computed using a 6-month lag.

<Insert Table V here.>

Returns on stocks in the lowest and highest liquidity quintile are 0.872% higher and 0.345% lower than the average 0.938% return, and the 1.217% spread between the lowest and highest liquidity quintile is statistically and economically significant. For S&P500 stocks, however, the 0.406% difference in return between the lowest and highest liquidity quintiles is statistically insignificant. The statistically significant negative signs on the interaction coefficients of S&P500 stock membership with liquidity quintile grouping, $I_{S\&P} \times PMLIQ_{i,q}$, can largely be attributed to lower returns on S&P500 stocks.

C. Information Efficiency

Improvements in liquidity from the primary activities of APs should reduce trading costs and enhance price discovery. One the one hand, if higher returns from $ETF\ Ownership$ are the result of a reduction in the supply of shares available for trade, then we should expect a decline in returns from the primary market activities of APs. On the other hand, if the increase in return is from an increase in undiversifiable risk, then we should expect a higher return.

In Glosten et al. (2019), the change in $ETF\ Ownership$ impounds information on systematic aggregate earnings across firms onto the current returns of ETF-owned stocks. We should not expect the primary market activities of APs to influence ETF-owned stock returns from factor informed trading. We should, however, expect the primary activities of APs to impound firm-specific information from public disclosures onto the current returns of ETF-owned stocks more efficiently.

ETF Ownership

Table VII reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over our sample period 2000 to 2019. In Panel A, we examine the interaction of liquidity from the primary market activities of APs with $ETF\ Ownership$, and in Panel B, the interaction with $\Delta ETF\ Ownership$.

<Insert Table VI here.>

Column 3 of Panel A confirms that higher returns from $ETF\ Ownership$ stem from illiquidity associated with share lockup. In the highest quintile of liquidity, there is a -0.027 (=-0.478/17.68) standard deviation decline in stock return associated with a one standard deviation increase in $ETF\ Ownership$. The decline represents 66.2% (=0.478/0.721) of the impact of $ETF\ Ownership$ on stock returns.

Column 3 of Panel B shows the primary market activities of APs have no significant effect on the incorporation of systematic market-wide information from factor informed trading on stock returns. The impact of the primary activities of APs on the liquidity of ETF-owned stocks will, however, impound information in public earnings announcements into current returns more efficiently.

Earnings Surprise

Using earnings announcements from the IBES database, we estimate Fama-MacBeth (1973) cross-sectional regressions each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \beta_1 \cdot PMLIQ_{i,q} + \beta_2 \cdot |\Delta EPS_t| + \beta_3\left(|\Delta EPS_t| \cdot PMLIQ_{i,q}\right)$$
$$+ \gamma_1\left(|\Delta EPS_t| \cdot PMLIQ_{i,q} \cdot I_t\right) + \gamma_2\left(|\Delta EPS_t| \cdot PMLIQ_{i,q} \cdot I_{t-1}\right) + \textstyle\sum_k \delta_k X_{ik} + \varepsilon_i \qquad (13)$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ_{i,q}$ are categorical variables that indicate the liquidity quintile rank to which stocks belong.

We align monthly stock returns with quarterly earnings announcement months as follows. If month t is not but month $t-1$ is a quarterly earnings announcement month, we assign the quarterly earnings announcement in month $t-1$ to month t. If month t is neither a quarterly earnings announcement month nor month after a quarterly earnings announcement but $t+1$ is a quarterly earnings announcement month, we assign the quarterly earnings announcement in month $t+1$ to month t. $|\Delta EPS_t|$ is a dummy variable which equals 1 when earnings announcements miss or beat median analyst forecasts in the subsequent quarter by at least 10%, and 0 otherwise. $I_t = 1$ if month t is an earnings announcement month and 0 otherwise. $I_{t-1} = 1$ if month is prior to an earnings announcement month and 0 otherwise. Control variables X_{ik} defined in Table I are demeaned and standardized each month. We exclude stocks with negative earnings announcements.

<Insert Table VII here.>

Table VII reports time-series averages of coefficients in Panel A when earnings announcements exceed analyst forecasts, and in Panel B, when earnings announcements are equal to or below analyst forecasts. Newey-West standard errors of the time-series average of estimated model parameters are computed using a 6-month lag.

The impact of the primary market activities of APs on liquidity are more economically significant when unexpected earnings surprises are negative. Contrasting the results in Columns 1 and 2 with 3 and 4, the spread of 1.304% between the highest and lowest quintile of liquidity when actual earnings miss forecast is almost twice the spread of 0.670% when actual earnings beat forecast.

When actual earnings beat analyst forecast, Column 1 shows returns on stocks in the lowest liquidity quintile are 0.514% higher, and 0.932% lower in the highest liquidity quintile, than average return. In Column 2, when $ETF\ Ownership$ and $\Delta ETF\ Ownership$ are also taken into account, returns on stocks in the lowest liquidity quintile are 0.240% higher and 0.430% lower in the highest liquidity quintile, than average return.

When actual earnings miss analyst forecast, Column 3 shows returns on stocks in the lowest liquidity quintile are 0.800% higher and 1.206% lower in the highest liquidity quintile, than average return. In Column 4, when $ETF\ Ownership$ and $\Delta ETF\ Ownership$ are also taken into account, returns on stocks in the lowest liquidity quintile are 0.538% higher and 0.766% lower in the highest liquidity quintile, than average return.

The impact of the primary market activities of APs on liquidity attenuate stock price declines associated with negative earnings surprises. Contrasting the coefficients on the interaction of liquidity with unexpected earnings surprises in Columns 2 and 4, liquidity is economically and statistically significant only when actual earnings miss analyst forecast. Returns will be lower by 0.233% in the lowest liquidity quintile and higher by 0.370% in the highest liquidity quintile than the average 0.578% decline.

Attenuating overreactions to negative earnings surprises reduce post earnings drifts. Finding supports the view that ETF activities enhance information efficiency.

VI. Asset and Stock Return Volatility

Stocks are embedded real options. Equity value is a convex function of the value of the underlying assets of the firm. Stock return and volatility will be positively correlated. At the same time, the unobservability of returns on the underlying assets of firms creates information uncertainty that deter ambiguity averse investors. Idiosyncratic risk will be priced. The positive relation between stock return and volatility will be stronger when improvements in liquidity from the primary market activities of APs enables new information to be impounded into the prices of the underlying stocks more quickly.

Fama-MacBeth (1973) cross-sectional regressions are estimated each month over our sample period. The dependent variable ret_i are monthly stock returns expressed in percent.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \gamma_1 IU_i + \gamma_2 \sum_{q=Low}^{High} \left(IU_i \cdot PMLIQ_{i,q}\right) + \sum_k \delta_k X_{ik} + \varepsilon_i \quad (14)$$

$PMLIQ_{i,q}$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest liquidity quintiles, and 0 otherwise. Proxy for information uncertainty IU_i and control variables X_{ik} are defined in Table I. All continuous variables are demeaned and standardized each month. $IU_i \cdot PMLIQ_{i,q}$ is the interaction of information uncertainty with the impact of the primary market activities of APs on liquidity. Time-series averages of the estimated coefficients are reported in the tables. Newey-West autocorrelated corrected standard errors assuming a 6-month lag are used to test for statistical significance. Results when we use changes in the proxies for information uncertainty instead are unchanged.

A. Convexity

Asset Return Volatility

In Column 2, returns on stocks in the lowest and highest liquidity quintile are 0.514% higher and 0.719% lower than the average 0.889% return, and the 1.233% spread between the lowest and highest liquidity quintile is statistically and economically significant.

<Insert Table VIII here.>

As expected, stock return is positively related to asset return volatility when stocks are embedded real options. A one percent change in asset return volatility will change stock return by 0.188% (=0.229/0.280). Furthermore, the impact of the primary market activities of APs on liquidity strengthens the stock return-volatility relation. Stock return-volatility relation is unchanged in the lowest liquidity quintile, but stronger in the highest liquidity quintile. In the highest liquidity quintile, a one percent change in asset return volatility will change stock return further by 1.125% (=0.315/0.280). The primary market activities of APs do not increase asset return volatility. Rather, that stock returns become more sensitive to volatility.

Stock Return Volatility

In Column 2, returns on stocks in the lowest and highest liquidity quintile are 0.437% higher and 1.097% lower than the average 1.045% return, and the 1.534% spread between the lowest and highest quintile is statistically and economically significant.

<Insert Table IX here.>

Stock returns are also positively related to equity return volatility when stocks are embedded real options. A one percent change in equity return volatility will change stock return by 0.090% (=1.589/17.68). Again, the impact of the primary market activities of APs on liquidity strengthens the stock return-volatility relation. Stock returns are unaffected by equity return volatility in the lowest liquidity quintile, but in the highest liquidity quintile, a one percent change in equity return volatility will change stock return further by 0.619% (=1.218/17.68). The primary market activities of APs make stock returns more sensitive to volatility.

B. Information Uncertainty

Idiosyncratic Stock Return Volatility

In Column 2, returns on stocks in the lowest and highest liquidity quintile are 0.456% higher and 0.981% lower than the average 1.064% return, and the 1.437% spread between the lowest and highest liquidity quintile is statistically and economically significant.

<Insert Table X here.>

As conjectured, idiosyncratic return volatility is priced when information uncertainty deters ambiguity averse investors. Stock returns increase with idiosyncratic return volatility. A one standard deviation increase in idiosyncratic return volatility will result in a 0.094 (=1.663/17.68) standard deviation increase in stock return. The primary market activities of APs on liquidity intensify the adverse effect of information uncertainty on participation by ambiguity averse investors. In the highest liquidity quintile, a one standard deviation increase in idiosyncratic stock return volatility results in a 0.039 (=0.687/17.68) standard deviation increase in stock return, but a 0.012 (=-0.220/17.68) standard deviation decrease in stock return in the lowest liquidity quintile.

Standard Deviation of Turnover

In Column 2, returns on stocks in the lowest and highest liquidity quintile are 0.504% higher and 0.833% lower than the average 1.024% return, and the 1.338% spread between the lowest and highest liquidity quintile is statistically and economically significant.

<Insert Table XI here.>

The standard deviation in turnover is another proxy for idiosyncratic risk. We find that stock returns increase with turnover volatility on ETF-owned stocks. A one standard deviation increase in turnover volatility will result in a 0.094 (=1.081/17.68) standard deviation increase in stock return. The primary market of activities of APs on liquidity intensify the adverse selection effect of information uncertainty. A one standard deviation increase in stock turnover volatility results in a 0.027 (=0.479/17.68) standard deviation increase in stock return in the highest liquidity quintile but a 0.011 (=-0.199/17.68) standard deviation decrease in stock return in the lowest liquidity quintile.

VII. Concluding Remarks

SPDR S&P 500 ETF Trust (SPY), first introduced in 1993 by State Street Global Advisors, is the oldest and largest U.S. listed and domiciled equity ETF as of 2019. Between 1993 and August 2019, assets under management (AUM) by index funds which track broad US equity indexes grew to $4.27 trillion, compared to only $4.25 trillion in U.S. listed actively managed equity funds.[19] The value of ETF shares traded is roughly 28% of the aggregate value of shares traded in U.S. exchanges (Boroujerdi and Fogertey, 2015; and Pisani, 2015).

Theory suggests that ETFs represent an important innovation. First, composite securities like ETFs are not redundant when uninformed investors have to trade to meet immediate liquidity needs, but prices are not fully revealing when some investors are informed. To avoid trading against informed

[19]WSJ "Where ETFs are headed in 2019" reports that $295 billion flowed into US domiciled ETFs in 2018 alone; 66.8% into stock funds and the remainder to fixed-income funds. 0.3% flowed out of alternative investment funds. WSJ "Index Funds Are the New Kings of Wall Street" reports that as of August 2019, assets under management in index equity funds with $4.27 trillion exceed actively managed equity funds with $4.25 trillion. https://www.wsj.com/articles/index-funds-are-the-new-kings-of-wall-street-11568799004?mod=searchresults&page=1&pos=4.

investors, uninformed investors will choose to meet their liquidity needs through ETF shares rather than individual stock ownership (Gorton and Pennachi, 1993; and Subrahmanyam, 1991). For uninformed investors, ETFs expand investment opportunities when transactions and holding costs would otherwise limit stock ownership to transparent and liquid stocks. Second, as Cong and Xu (2019) and Glosten et al. (2019) point out, ETFs are well-diversified portfolios whose returns largely reflect systematic market rather than idiosyncratic firm-specific factors. The low cost and intraday trading of ETFs make informed market factor investment strategies by sophisticated asset managers feasible.

Herein, we use fund flow shocks from exogenous changes in the demand for ETF shares to quantify the volume of stocks purchased or sold by APs in conjunction with the creation or redemption of ETF shares. We introduce a novel measure of the liquidity impact of the primary market activities of APs from the projected cost of trading the underlying stocks. We find the impact of the primary market activities on liquidity is priced onto the returns on stocks underlying ETFs. Returns on ETF-owned stocks are lower when liquidity is high. The negative relation between return and liquidity is a stock-specific effect distinct from and unexplained by systematic market, size, book-to-market, momentum, investment, or profitability asset pricing factors. Further, we find the improvements in liquidity from the primary market activities of APs enhances information efficiency and strengthens the stock return-volatility relation.

Our empirical findings are consistent with the experience of U.S. ETFs during the COVID-19 crisis documented by the SEC.[20] Despite unprecedented market volatility in March 2020 caused by the COVID-19 crisis, APs stepped up and facilitated a significantly higher level of creations and redemptions of ETF shares in March 2020 than during a comparable "normal" period in March 2019. ETF shares traded smoothly and efficiently on the stock exchanges and acted as a price discovery tool

[20]See https://www.sec.gov/comments/credit-market-interconnectedness/cll10-2.pdf.

for investors. Contrary to predictions by some policymakers and other observers, the ETF ecosystem remained strong and functioned well.

References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Anton, M., & Polk, C. (2014). Connected Stocks. *The Journal of Finance*, *69*(3), 1099–1127.

Bae, K., & Kim, D. (2020). Liquidity risk and exchange-traded fund returns, variances, and tracking errors. *Journal of Financial Economics*, *138*(1), 222–253.

Basak, S., & Pavlova, A. (2013). Asset Prices and Institutional Investors. *American Economic Review*, *103*(5), 1728–1758.

Ben-David, I., Franzoni, F., & Moussawi, R. (2018). Do ETFs Increase Volatility? *The Journal of Finance*, *73*(6), 2471–2535.

Berk, J. B., Green, R. C., & Naik, V. (1999). Optimal Investment, Growth Options, and Security Returns. *The Journal of Finance*, *54*(5), 1553–1607.

Bharath, S. T., & Shumway, T. (2008). Forecasting Default with the Merton Distance to Default Model. *The Review of Financial Studies*, *21*(3), 1339–1369.

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*(3), 637–654.

Boroujerdi, R., & Fogertey, K. (2015, April 10). ETFs: The Rise of the Machines. *Goldman Sachs Equity Research*.

Box, T., Davis, R., Evans, R., & Lynch, A. (2021). Intraday arbitrage between ETFs and their underlying portfolios. *Journal of Financial Economics*, *141*(3), 1078–1095.

Brockman, P., Chung, D. Y., & Pérignon, C. (2009). Commonality in Liquidity: A Global Perspective. *The Journal of Financial and Quantitative Analysis*, *44*(4), 851–882.

Broman, M. S., & Shum, P. (2018). Relative Liquidity, Fund Flows and Short-Term Demand: Evidence from Exchange-Traded Funds. *Financial Review*, *53*(1), 87–115.

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82.

Cespa, G., & Foucault, T. (2014). Illiquidity Contagion and Liquidity Crashes. *The Review of Financial Studies*, *27*(6), 1615–1660.

Chordia, T., Subrahmanyam, A., & Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, *59*(1), 3–32.

Cong, L. W., & Xu, D. (2019). Rise of Factor Investing: Asset Prices, Informational Efficiency, and Security Design. *SSRN*.

Da, Z., & Shive, S. (2018). Exchange traded funds and asset return correlations. *European Financial Management*, *24*(1), 136–168.

Daniel, K., & Titman, S. (2006). Market Reactions to Tangible and Intangible Information. *The Journal of Finance*, *61*(4), 1605–1643.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance*, *45*(2), 379–395.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., & Peijnenburg, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, *119*(3), 559–577.

Duffee, G. R. (1995). Stock returns and volatility a firm-level analysis. *Journal of Financial Economics*, *37*(3), 399–420.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, *75*(4), 643–669.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fama, E. F., & MacBeth, J. D. (1973a). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636.

Fama, E. F., & MacBeth, J. D. (1973b). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636.

Franzoni, F., & Schmalz, M. C. (2017). Fund Flows and Market States. *The Review of Financial Studies*, *30*(8), 2621–2673.

French, K. R. (2008). Presidential Address: The Cost of Active Investing. *The Journal of Finance*, *63*(4), 1537–1573.

Glosten, L., Nallareddy, S., & Zou, Y. (2020). ETF Activity and Informational Efficiency of Underlying Securities. *Management Science*, *67*(1), 22–47.

Goldstein, I., Jiang, H., & Ng, D. T. (2017). Investor flows and fragility in corporate bond funds. *Journal of Financial Economics*, *126*(3), 592–613.

Gorton, G. B., & Pennacchi, G. G. (1993). Security Baskets and Index-Linked Securities. *The Journal of Business*, *66*(1), 1–27.

Greenwood, R., & Thesmar, D. (2011). Stock price fragility. *Journal of Financial Economics*, *102*(3), 471–490.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, *4*(1), 5–28.

Huang, J., Wei, K. D., & Yan, H. (2007). Participation Costs and the Sensitivity of Fund Flows to Past Performance. *The Journal of Finance*, *62*(3), 1273–1311.

Israeli, D., Lee, C. M. C., & Sridharan, S. A. (2017). Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies*, *22*(3), 1048–1083.

Iwadate, Bruce, Mind Thy Neighbor's Portfolio: A Network Approach to Contagion in the ETF Market (April 1, 2021). Available at SSRN: https://ssrn.com/abstract=3535942

Jiang, G., Lee, C. M. C., & Zhang, Y. (2005). Information Uncertainty and Expected Returns. *Review of Accounting Studies*, *10*(2), 185–221.

Keim, D. B., & Madhavan, A. (1997). Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics*, *46*(3), 265–292.

Lee, C. M. C., & Swaminathan, B. (2000). Price Momentum and Trading Volume. *The Journal of Finance*, *55*(5), 2017–2069.

Lou, D. (2012). A Flow-Based Explanation for Return Predictability. *The Review of Financial Studies*, *25*(12), 3457–3489.

Lynch, A. W., & Musto, D. K. (2003). How Investors Interpret Past Fund Returns. *The Journal of Finance*, *58*(5), 2033–2058.

McDonald, R., & Siegel, D. (1986). The Value of Waiting to Invest. *The Quarterly Journal of Economics*, *101*(4), 707–728.

Merton, R. C. (1974). ON THE PRICING OF CORPORATE DEBT: THE RISK STRUCTURE OF INTEREST RATES*. *The Journal of Finance*, *29*(2), 449–470.

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, *42*(3), 483–510.

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

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2020). Fund tradeoffs. *Journal of Financial Economics*, *138*(3), 614–634.

Pisani, R. (2015, July 2). Here is why ETFs are a growing part of total trading value. *CNBC*. Here is why ETFs are a growing part of total trading value

Subrahmanyam, A. (1991). A Theory of Trading in Stock Index Futures. *The Review of Financial Studies*, *4*(1), 17–51.

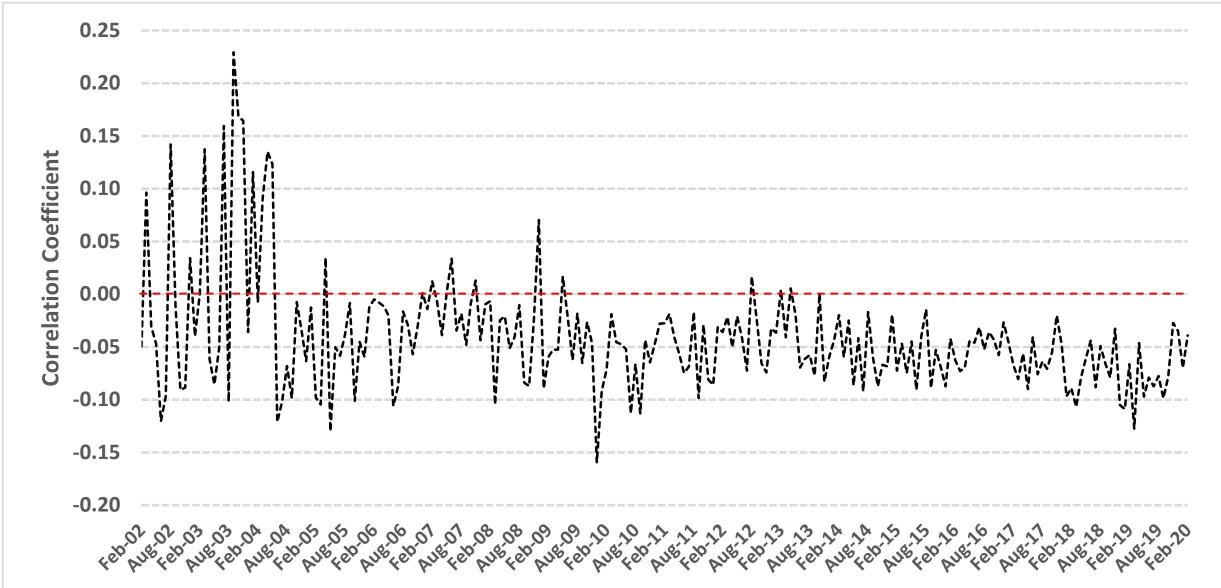


Figure I: Amihud Illiquidity and Fund Flow-Driven Trading Cost

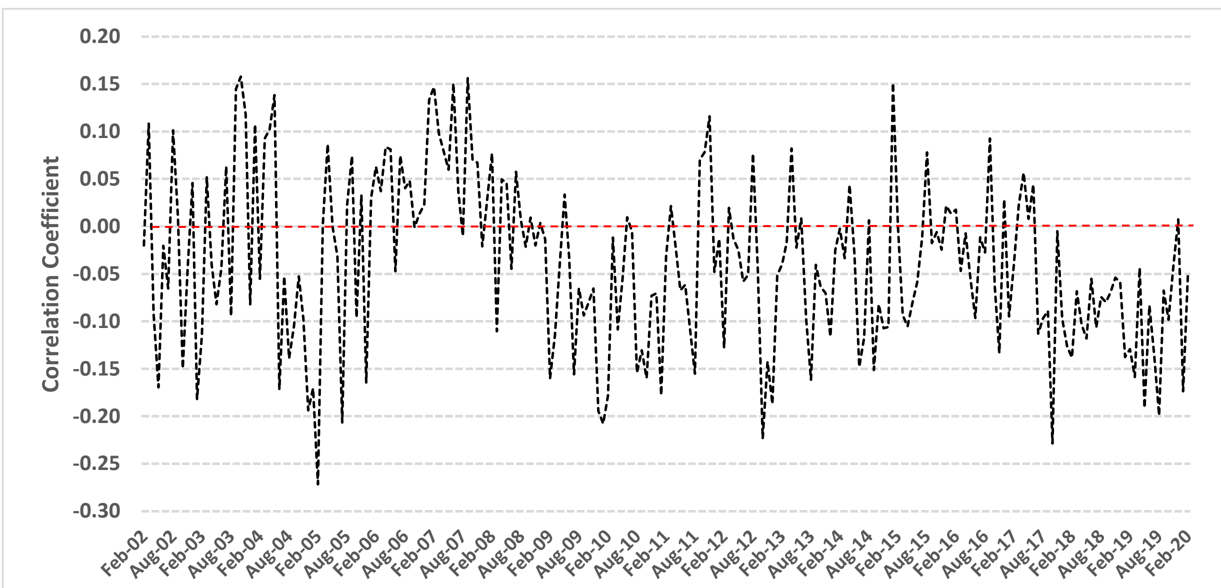


Figure II: Stock Return Volatility and Fund Flow-Driven Trading Cost
Fund flow-driven trading cost is computed as the product of aggregate fund ownership, ownership concentration, and ownership concentration weighted volatility in fund flow. Amihud (2002) is the ratio of daily absolute return to daily dollar volume of shares traded averaged over days in a month. Stock return volatility is the standard deviation of daily returns in the past 25 days.

Table I: Summary Statistics

Table reports summary statistics on variables used in this study. Sample period is January 2002 through December 2019. Number of observations is 792,696. $PMLIQ_i$ denotes the liquidity provided by the primary market activity of APs. ETF ownership is the total number of shares owned by ETFs at month end as a percentage of total shares outstanding. Delta ETF ownership is the month-by-month change in ETF ownership for a given stock. Market capitalization is the market value of equity computed as closing price multiplied by total shares outstanding. Book-to-market is book equity to shareholders' equity calculated following Daniel and Titman (2006). Momentum is the cumulative monthly return in the preceding 12-month period, ending in 1 month prior to start of the month. Market beta is the coefficient on market return premium estimated with CAPM model using rolling 36 months returns. Amihud illiquidity $illiq_{i,d} = \sum_d |r_{i,d}|/dvol_{i,d}$ is absolute daily return divided by daily trading volume in millions of dollars averaged across days in the month. Turnover is French (2008) adjusted CRSP volume divided by shares outstanding. Stock returns are expressed in percent. Asset volatility is estimated following Bharath and Shumway (2008) which captures the volatility of asset value for a given firm. Stock return volatility is estimated as the standard deviation of daily returns of past 25 trading days. Idiosyncratic volatility is rolling 36-month window standard deviation of monthly stock return residuals, which is estimated from CAPM model using rolling 36 months returns. Turnover standard deviation is the standard deviation of daily stock turnover over a month, with daily turnover estimated as daily trading volume scaled by shares outstanding. We take natural log of turnover standard deviation and get $Ln(Turnover\ Standard\ Deviation)$.

	Mean	**Std Dev**	**P25**	**Median**	**P75**
$Stock\ Return$ (%)	0.873	17.684	-6.434	0.397	7.027
$PMLIQ \cdot 10^{-4}$	5.760	34.320	0.217	0.376	0.741
$ETF\ Ownership$	0.039	0.039	0.009	0.026	0.059
$\Delta ETF\ Ownership \cdot 10^3$	0.405	5.267	-0.117	0.024	0.854
$ln(mktcap)$	20.164	1.953	18.800	20.072	21.452
$ln(book/market)$	0.725	0.854	0.165	0.639	1.209
$Momentum$	0.111	0.569	-0.216	0.052	0.319
$Market\ Beta$	1.259	1.017	0.588	1.099	1.726
$ln(Amihud\ Illiquidity)$	-7.062	3.105	-9.369	-7.331	-4.996
$Asset\ Volatility$	0.434	0.275	0.248	0.353	0.532
$Stock\ Return\ Volatility$	0.030	0.022	0.016	0.024	0.037
$Idio\_Stock\ Return\ Volatility$	0.126	0.079	0.070	0.105	0.159
$Std\_Turnover$	0.011	0.015	0.003	0.007	0.012
$ln(Std\_Turnover)$	-5.016	0.955	-5.656	-5.035	-4.419

Table II: Correlation Table

Table reports the correlation between variables used in this study. Variable definitions can be found in Table I.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	*Stock Return*	1													
2	*PMLIQ*	-0.0016	1												
3	*ETF Ownership*	-0.0011	-0.2098	1											
4	*ΔETF Ownership*	0.0223	-0.0215	0.128	1										
5	*ln(mktcap)*	0.0447	-0.1678	0.3992	0.0257	1									
6	*ln(book/market)*	0.0818	-0.0381	0.0489	0.0189	0.2633	1								
7	*Momentum*	-0.0076	-0.0167	0.0342	0.0348	0.1655	0.2415	1							
8	*Market Beta*	0.0201	-0.0331	0.0595	0.0017	-0.0392	0.067	-0.0193	1						
9	*ln(Amihud Illiquidity)*	-0.0195	0.2001	-0.5125	-0.0353	-0.9374	-0.273	-0.168	-0.0451	1					
10	*Asset Volatility*	0.0007	0.0823	-0.2551	-0.0123	-0.4338	0.1097	-0.0594	0.3204	0.3482	1				
11	*Stock Return Volatility*	0.0643	0.0783	-0.2137	-0.0132	-0.4359	-0.006	-0.118	0.3012	0.3693	0.6169	1			
12	*Idio_Stock Return Volatility*	0.0369	0.0622	-0.2354	-0.0105	-0.4491	0.0786	0.0144	0.4337	0.3374	0.6797	0.6003	1		
13	*Std_Turnover*	0.0003	0.0034	-0.0085	0.0054	-0.1388	0.1026	-0.0318	0.2724	-0.0146	0.4631	0.325	0.5154	1	
14	*ln(Std_Turnover)*	0.0002	-0.0516	0.1325	0.0164	-0.062	0.1422	-0.0189	0.3536	-0.1473	0.4411	0.3297	0.5096	0.7983	1

Table III: Liquidity Impact of the Primary Market Activity of Authorized Participants

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i, q$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard error of the time-series average of the difference between high and low quintiles are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

	DV=Stock Return				
	1	2	3	4	5
$PMLIQ$: $High - Low\ Quintile$	-0.152	-1.925***	-1.303***		
	(0.444)	(0.000)	(0.000)		
N	216	216	216		
$PMLIQ$: $Low\ Quintile$	0.196*	0.857***	0.623***		
	(0.053)	(0.000)	(0.000)		
$PMLIQ$: $High\ Quintile$	0.045	-1.068***	-0.680***		
	(0.777)	(0.000)	(0.000)		
$PMLIQ$				-0.152***	-0.142
				(0.000)	(0.328)
$ETF\ Ownership\_lag$			0.511***	0.748***	0.764***
			(0.000)	(0.000)	(0.000)
$PMLIQ \times ETF\ Ownershp\_lag$					0.164
					(0.113)
$ln(mktcap)$		6.842***	7.048***	6.824***	6.858***
		(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$		1.335***	1.368***	1.370***	1.376***
		(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$		-0.461***	-0.438***	-0.394***	-0.413***
		(0.000)	(0.000)	(0.000)	(0.000)
$Market\ Beta$		0.649**	0.661**	0.669**	0.679**
		(0.043)	(0.038)	(0.036)	(0.034)
$ln(Amihud\ Illiquidity)\_lag$		6.869***	7.258***	7.187***	7.203***
		(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.913**	0.987**	0.950**	0.940**	0.959**
	(0.037)	(0.028)	(0.034)	(0.034)	(0.032)
N	785852	674113	674113	670948	665398
$adjusted\ R^2$	0.004	0.115	0.117	0.115	0.116

	DV=One Month Forward Stock Return				
	1	2	3	4	5
$PMLIQ: High - Low\ Quintile$	0.274	0.102	0.203		
	(0.214)	(0.681)	(0.286)		
N	216	216	216		
$PMLIQ: Low\ Quintile$	-0.113	0.043	0.017		
	(0.255)	(0.664)	(0.852)		
$PMLIQ: High\ Quintile$	0.161	0.145	0.220*		
	(0.356)	(0.379)	(0.094)		
$PMLIQ$				-0.011	-0.185*
				(0.745)	(0.086)
$ETF\ Ownership\_lag$			0.105	0.068	0.037
			(0.179)	(0.476)	(0.706)
$PMLIQ \times ETF\ Ownershp\_lag$					-0.207**
					(0.027)
$ln(mktcap)$		-0.586*	-0.584	-0.508	-0.544
		(0.075)	(0.103)	(0.135)	(0.113)
$ln(book/market)$		1.350***	1.357***	1.349***	1.348***
		(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$		-0.407***	-0.398***	-0.411***	-0.409***
		(0.001)	(0.002)	(0.002)	(0.002)
$Market\ Beta$		-0.061	-0.056	-0.051	-0.029
		(0.681)	(0.705)	(0.730)	(0.831)
$ln(Amihud\ Illiquidity)\_lag$		-0.104	-0.040	-0.000	-0.033
		(0.696)	(0.900)	(0.999)	(0.914)
$Constant$	0.924**	1.001**	0.983**	1.019**	1.018**
	(0.035)	(0.020)	(0.023)	(0.022)	(0.021)
N	782725	673736	673736	670575	665027
$adjusted\ R^2$	0.004	0.049	0.050	0.049	0.049

Table IV: Asset Pricing Factors and Primary Market Liquidity

Table reports time-series regressions over a sample period 2002 to 2019. Dependent variable $ret_t(PMLIQ\_q)$ is the time-series of estimated returns from monthly cross-sectional regressions of stock returns on $PMLIQ\_i,q$ dummy variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise.

$$ret_t(PMLIQ\_q) = \alpha + \sum_k \delta_k X_{k,t} + \varepsilon_t$$

X_{ik} denote control variables that are demeaned and standardized each month. Control variables $X_{k,t}$ are asset pricing factors defined in Table I are demeaned and standardized each month. Newey-West standard errors of estimated model parameters are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

	DV = Return on Stocks in $PMLIQ$: $Low\ Quintle$					**DV = Return on Stocks in $PMLIQ$: $High\ Quintle$**				
	1	2	3	4	5	1	2	3	4	5
Constant	0.623***	0.627***	0.639***	0.644***	0.599***	-0.680***	-0.666***	-0.668***	-0.630***	-0.695***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
MKT-RF		-0.006	-0.035	-0.040	0.002		-0.020	0.016	-0.027	-0.008
		(0.781)	(0.152)	(0.180)	(0.954)		(0.566)	(0.676)	(0.428)	(0.824)
SMB			0.055	0.057				-0.136**	-0.122**	
			(0.251)	(0.237)				(0.013)	(0.026)	
HML			0.155***	0.149***				-0.064	-0.111	
			(0.009)	(0.005)				(0.320)	(0.138)	
MOM				-1.146					-9.679**	
				(0.630)					(0.048)	
CMA					0.131**					0.050
					(0.039)					(0.673)
RMW					0.028					0.051
					(0.758)					(0.670)
N	216	216	216	216	216	216	216	216	216	216
R^2	0.000	-0.004	0.070	0.067	0.006	0.000	-0.003	0.027	0.063	-0.009

Table V: S&P500 Membership and Primary Market Liquidity

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + I_{S\&P} \cdot \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i, q$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise, and $I_{S\&P}$ is equal to 1 if the stock is in the S&P500, and 0 otherwise. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard errors of the time series average of estimated model parameters are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

	DV = Stock Return				
	1	2	3	4	5
$PMLIQ$: $High - Low\ Quintile$	-0.152	-1.964***			
	(0.469)	(0.000)			
$S\&P500$: $PMLIQ$: $High - Low\ Quintile$	-0.342	0.477			
	(0.305)	(0.112)			
N	216	216			
$PMLIQ$: $Low\ Quintile$	0.241**	1.218***			
	(0.015)	(0.000)			
$PMLIQ$: $High\ Quintile$	0.089	-0.746***			
	(0.631)	(0.000)			
$S\&P500 \times PMLIQ$: $Low\ Quintile$	0.193	-1.973***			
	(0.456)	(0.000)			
$S\&P500 \times PMLIQ$: $High\ Quintile$	-0.150	-1.496***			
	(0.555)	(0.000)			
$PMLIQ$			-0.328**	-0.233**	-0.061
			(0.030)	(0.016)	(0.669)
$S\&P500 \times PMLIQ$			27.363***	24.159***	24.177***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership\_lag$				0.464***	0.492***
				(0.000)	(0.000)
$PMLIQ \times ETF\ Ownership\_lag$					0.214**
					(0.036)
$ln(mktcap)$		7.079***	7.589***	7.727***	7.769***
		(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$		1.326***	1.267***	1.309***	1.314***
		(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$		-0.492***	-0.563***	-0.530***	-0.550***
		(0.000)	(0.000)	(0.000)	(0.000)
$Market\ Beta$		0.648**	0.626*	0.642**	0.652**
		(0.044)	(0.053)	(0.045)	(0.043)
$ln(Amihud\ Illiquidity)\_lag$		6.905***	6.779***	7.230***	7.254***
		(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.913**	0.971**	1.376***	1.318***	1.360***
	(0.037)	(0.031)	(0.002)	(0.003)	(0.002)
N	785852	674113	670948	670948	665398
$adjusted\ R^2$	0.005	0.116	0.117	0.119	0.120

Table VI: ETF Ownership and Primary Market Liquidity

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \sum_{q=Low}^{High} \beta_q \cdot Own_{i,q} \cdot PMLIQ_{i,q} + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i,q$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. In Panel A, $Own\_i,q = ETF\ Ownership\_lag$ and in Panel B, $Own\_i,q = \Delta ETF\ Ownership$. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard errors of the time series average of estimated model parameters are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

	DV = Stock Return				
	1	2	3	4	5
$PMLIQ: High - Low\ Quintile$	-0.287	-1.157***			
	(0.203)	(0.000)			
$ETF\ Ownership\_lag: PMLIQ: High - Low\ Quintile$	0.0604	0.369*			
	(0.780)	(0.0525)			
N	216	216			
$PMLIQ: Low\ Quintile$	0.311**	0.725***			
	(0.011)	(0.000)			
$PMLIQ: High\ Quintile$	-0.146	-1.082***			
	(0.479)	(0.000)			
$ETF\ Ownership\_lag \times PMLIQ: Low\ Quintile$	-0.216	0.027			
	(0.348)	(0.879)			
$ETF\ Ownership\_lag \times PMLIQ: High\ Quintile$	-0.128	0.219***			
	(0.128)	(0.002)			
$PMLIQ$			-0.152***	0.127	-0.142
			(0.000)	(0.287)	(0.328)
$ETF\ Ownership\_lag$			0.748***	-0.020	0.764***
			(0.000)	(0.803)	(0.000)
$PMLIQ \times ETF\ Ownership\_lag$				0.108	0.164
				(0.348)	(0.113)
$ln(mktcap)$		6.929***	6.824***		6.858***
		(0.000)	(0.000)		(0.000)
$ln(book/market)$		1.342***	1.370***		1.376***
		(0.000)	(0.000)		(0.000)
$Momentum$		-0.459***	-0.394***		-0.413***
		(0.000)	(0.000)		(0.000)
$Market\ Beta$		0.653**	0.669**		0.679**
		(0.042)	(0.036)		(0.034)
$ln(Amihud\ Illiquidity)\_lag$		6.950***	7.187***		7.203***
		(0.000)	(0.000)		(0.000)
$Constant$	0.881**	0.987**	0.940**	0.973**	0.959**
	(0.045)	(0.028)	(0.034)	(0.034)	(0.032)
N	778829	670948	670948	770232	665398
$adjusted\ R^2$	0.006	0.116	0.115	0.005	0.116

Table VII: Earnings Surprise and Primary Market Liquidity

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \beta_1 \cdot PMLIQ_{i,q} + \beta_2 \cdot |\Delta EPS_t| + \beta_3\left(|\Delta EPS_t| \cdot PMLIQ_{i,q}\right) + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i,q$ are categorical variables that indicate the quintile rank to which stocks belong. We align monthly stock returns with quarterly earnings announcement months as follows. If month t is not but month $t-1$ is a quarterly earnings announcement month, we assign the quarterly earnings announcement in month $t-1$ to month t . If month t is neither a quarterly earnings announcement month nor month after a quarterly earnings announcement but $t+1$ is a quarterly earnings announcement month, we assign the quarterly earnings announcement in month $t+1$ to month t. $|\Delta EPS_t|$ is a dummy variable which equals 1 when earnings announcements miss or beat median analyst forecasts in the subsequent quarter by at least 10%, and 0 otherwise. Control variables X_{ik} defined in Table I are demeaned and standardized each month. We exclude stocks with negative earnings announcements. Panel A reports results when earnings announcements exceed analyst forecasts, and in Panel B, when earnings announcements are equal to or below analyst forecasts. Newey-West standard errors of the time-series average of estimated model parameters are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

	Actual EPS >= Analyst Forecast			**Actual EPS < Analyst Forecast**		
	1	2	3	4	5	6
$PMLIQ$: $Low\ Quintile$	0.514***			0.800***		
	(0.000)			(0.000)		
$PMLIQ$: $High\ Quintile$	-0.932***			-1.206***		
	(0.000)			(0.000)		
$PMLIQ$: $Low\ Quintile \times \Delta EPS$	0.171*			-0.213		
	(0.060)			(0.143)		
$PMLIQ$: $High\ Quintile \times \Delta EPS$	0.123			0.326*		
	(0.415)			(0.067)		
ΔEPS	1.752***			-0.614***		
	(0.000)			(0.000)		
$PMLIQ$		2.976***	2.868***		-2.859**	-2.934***
		(0.000)	(0.000)		(0.013)	(0.009)
$PMLIQ \times \Delta EPS$		-5.534***	-5.355***		3.201***	3.335***
		(0.000)	(0.000)		(0.000)	(0.000)
$\Delta ETF\ Ownership \times \Delta EPS$			0.488***			0.392***
			(0.000)			(0.000)
$ln(mktcap)$	5.238***	4.648***	4.682***	5.636***	5.153***	5.163***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.191***	1.138***	1.132***	0.926***	0.909***	0.908***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.442***	-0.376***	-0.380***	-0.848***	-0.803***	-0.810***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Market\ Beta$	0.435	0.509*	0.511*	0.145	0.101	0.108
	(0.121)	(0.074)	(0.074)	(0.564)	(0.694)	(0.675)
$ln(Amihud\ Illiquidity)\_lag$	6.191***	5.815***	5.851***	5.765***	5.300***	5.306***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	1.394***	2.427***	2.412***	0.023	-0.414	-0.427
	(0.001)	(0.000)	(0.000)	(0.955)	(0.327)	(0.313)

Table VII (Cont.)

	Actual EPS >= Analyst Forecast			Actual EPS < Analyst Forecast		
	1	2	3	4	5	6
N	314123	313452	313452	111594	111285	111285
adjusted R^2	0.131	0.121	0.123	0.133	0.127	0.130

Table VIII: Asset Volatility

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \gamma_1 \sigma_i^{asset} + \gamma_2 \sum_{q=Low}^{High} \left(\sigma_i^{asset} \cdot PMLIQ_{i,q}\right) + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i,q$ are indicator variables that equal for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. The volatility of return on the latent underlying assets of a firm, σ_i^{asset}, is estimated following Bharath and Shumway (2008). Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard error of the time-series average of the difference between high and low quintiles are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

Panel A: Asset Volatility	**DV = Stock Return**				
	1	2	3	4	5
$PMLIQ$: $Low\ Quintile$	0.914***				
	(0.000)				
$PMLIQ$: $High\ Quintile$	-1.243***				
	(0.000)				
$Asset\ Volatility \times PMLIQ$: $Low\ Quintile$	0.155				
	(0.208)				
$Asset\ Volatility \times PMLIQ$: $High\ Quintile$	0.457***				
	(0.001)				
$Asset\ Volatility$		0.162*	0.308***	0.238**	0.238**
		(0.095)	(0.002)	(0.013)	(0.014)
$PMLIQ$		-0.289***	-0.182***	-0.178***	-0.273*
		(0.000)	(0.000)	(0.000)	(0.057)
$PMLIQ \times Asset\ Volatility$		0.038	0.019	-0.040	-0.027
		(0.304)	(0.617)	(0.284)	(0.593)
$PMLIQ \times ETF\ Ownership_{lag}$					0.111
					(0.337)
$ETF\ Ownership_{lag}$			0.853***	0.857***	0.859***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times Asset\ Volatility$				-0.196***	-0.221***
				(0.005)	(0.002)
$ln(mktcap)$	6.601***	6.029***	6.707***	6.730***	6.759***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.192***	1.168***	1.205***	1.200***	1.205***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.381***	-0.345***	-0.316***	-0.309***	-0.324***
	(0.000)	(0.001)	(0.003)	(0.003)	(0.002)
$Market\ Beta$	0.597*	0.588*	0.570*	0.586*	0.600*
	(0.057)	(0.058)	(0.064)	(0.060)	(0.055)
$ln(Amihud\ Illiquidity)\_lag$	6.593***	5.990***	7.012***	7.029***	7.040***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.930**	0.918**	0.870*	0.830*	0.832*
	(0.039)	(0.039)	(0.051)	(0.062)	(0.063)
N	526225	523735	523735	523735	519488
$adjusted\ R^2$	0.122	0.119	0.124	0.126	0.127

Panel B: Δ Asset Volatility	DV = Stock Return				
	1	2	3	4	5
$PMLIQ: Low\ Quintile$	0.866***				
	(0.000)				
$PMLIQ: High\ Quintile$	-1.126***				
	(0.000)				
$\Delta Asset\ Volatility \times PMLIQ: Low\ Quintile$	0.314*				
	(0.057)				
$\Delta Asset\ Volatility \times PMLIQ: High\ Quintile$	1.501***				
	(0.000)				
$\Delta Asset\ Volatility$		0.609***	0.592***	0.387***	0.397***
		(0.000)	(0.000)	(0.007)	(0.006)
$PMLIQ$		-0.300***	-0.206***	-0.207***	-0.251
		(0.000)	(0.000)	(0.000)	(0.120)
$PMLIQ \times \Delta Asset\ Volatility$		0.169***	0.171***	-0.060	0.030
		(0.003)	(0.002)	(0.357)	(0.651)
$PMLIQ \times ETF\ Ownership_{lag}$					0.062
					(0.630)
$ETF\ Ownership_{lag}$			0.788***	0.781***	0.778***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times \Delta Asset\ Volatility$				-0.681***	-0.672***
				(0.000)	(0.000)
$ln(mktcap)$	6.335***	5.813***	6.350***	6.325***	6.347***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.225***	1.206***	1.265***	1.257***	1.263***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.368***	-0.324***	-0.299***	-0.297***	-0.312***
	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
$Market\ Beta$	0.642**	0.654**	0.673**	0.673**	0.686**
	(0.043)	(0.040)	(0.032)	(0.032)	(0.029)
$ln(Amihud\ Illiquidity)\_lag$	6.370***	5.821***	6.728***	6.695***	6.699***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.940**	0.923**	0.885**	0.883**	0.887**
	(0.038)	(0.039)	(0.048)	(0.047)	(0.049)
N	525058	522574	522574	522574	518435
$adjusted\ R^2$	0.128	0.124	0.129	0.133	0.134

Table IX: Stock Return Volatility

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \gamma_1 \sigma_i^{stock} + \gamma_2 \sum_{q=Low}^{High} \left(\sigma_i^{stock} \cdot PMLIQ_{i,q}\right) + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i,q$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. Stock return volatility, σ_i^{stock}, is computed following Jiang, Lee and Zhang (2005) as the standard deviation of daily returns over the past 25 trading days. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard error of the time-series average of the difference between high and low quintiles are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

Panel A: Stock Return Volatility			**DV = Stock Return**		
	1	2	3	4	5
$PMLIQ$: $Low\ Quintile$	0.972***				
	(0.000)				
$PMLIQ$: $High\ Quintile$	-1.616***				
	(0.000)				
$Stock\ Return\ Volatility \times PMLIQ$: $Low\ Quintile$	1.075***				
	(0.000)				
$Stock\ Return\ Volatility \times PMLIQ$: $High\ Quintile$	2.509***				
	(0.000)				
$Stock\ Return\ Volatility$		1.821***	1.896***	1.552***	1.559***
		(0.000)	(0.000)	(0.000)	(0.000)
$PMLIQ$		-0.365***	-0.261***	-0.259***	-0.246**
		(0.000)	(0.000)	(0.000)	(0.012)
$PMLIQ \times Stock\ Return\ Volatility$		0.129**	0.103**	-0.093**	0.017
		(0.026)	(0.045)	(0.028)	(0.700)
$PMLIQ \times ETF\ Ownership_{lag}$					0.191**
					(0.033)
$ETF\ Ownership_{lag}$			0.934***	0.845***	0.861***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times Stock\ Return\ Volatility$				-0.773***	-0.774***
				(0.000)	(0.000)
$ln(mktcap)$	7.318***	7.092***	7.651***	7.630***	7.658***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.187***	1.076***	1.144***	1.120***	1.126***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.370***	-0.325***	-0.291***	-0.270***	-0.286***
	(0.000)	(0.000)	(0.002)	(0.003)	(0.002)
$Market\ Beta$	0.475	0.235	0.238	0.294	0.307
	(0.109)	(0.366)	(0.354)	(0.261)	(0.240)
$ln(Amihud\ Illiquidity)\_lag$	6.998***	6.303***	7.316***	7.283***	7.298***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.936**	1.017**	0.991**	0.805*	0.835*
	(0.034)	(0.024)	(0.028)	(0.070)	(0.065)
N	674089	670924	670924	670924	665374
$adjusted\ R^2$	0.136	0.141	0.146	0.151	0.152

Panel A: Δ Stock Return Volatility	DV = Stock Return				
	1	2	3	4	5
$PMLIQ$: $Low\ Quintile$	0.846***				
	(0.000)				
$PMLIQ$: $High\ Quintile$	-1.082***				
	(0.000)				
$\Delta Stock\ Return\ Volatility \times PMLIQ$: $Low\ Quintile$	0.805***				
	(0.000)				
$\Delta Stock\ Return\ Volatility \times PMLIQ$: $High\ Quintile$	2.486***				
	(0.000)				
$\Delta Stock\ Return\ Volatility$		1.332***	1.326***	0.969***	0.982***
		(0.000)	(0.000)	(0.000)	(0.000)
$PMLIQ$		-0.305***	-0.219***	-0.216***	-0.246**
		(0.000)	(0.000)	(0.000)	(0.038)
$PMLIQ \times \Delta Stock\ Return\ Volatility$		0.258***	0.260***	0.025	0.145**
		(0.000)	(0.000)	(0.655)	(0.039)
$PMLIQ \times ETF\ Ownership_{lag}$					0.088
					(0.332)
$ETF\ Ownership_{lag}$			0.731***	0.713***	0.725***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times \Delta Stock\ Return\ Volatility$				-0.915***	-0.898***
				(0.000)	(0.000)
$ln(mktcap)$	6.702***	6.245***	6.664***	6.604***	6.633***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.318***	1.284***	1.344***	1.336***	1.339***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.418***	-0.382***	-0.357***	-0.342***	-0.359***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
$Market\ Beta$	0.646**	0.666**	0.681**	0.677**	0.686**
	(0.044)	(0.042)	(0.035)	(0.036)	(0.034)
$ln(Amihud\ Illiquidity)\_lag$	6.714***	6.216***	7.004***	6.928***	6.937***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.980**	0.960**	0.937**	0.936**	0.950**
	(0.029)	(0.030)	(0.035)	(0.035)	(0.034)
N	674088	670923	670923	670923	665373
$adjusted\ R^2$	0.134	0.133	0.137	0.143	0.143

Table X: Idiosyncratic Stock Return Volatility

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.
$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \gamma_1 idio\_\sigma_i^{stock} + \gamma_2 \sum_{q=Low}^{High} \left(idio\_\sigma_i^{stock} \cdot PMLIQ_{i,q}\right) + \sum_k \delta_k X_{ik} + \varepsilon_i$ Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i, q$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. Idiosyncratic volatility, $idio\_\sigma_i^{stock}$, is the standard deviation of the residuals from an estimated CAPM model using rolling 36-month windows. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard error of the time-series average of the difference between high and low quintiles are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

Panel A: Idiosyncratic Volatility	**DV = Stock Return**				
	1	2	3	4	5
$PMLIQ$: $Low\ Quintile$	1.024***				
	(0.000)				
$PMLIQ$: $High\ Quintile$	-1.513***				
	(0.000)				
$Idio\_Volatility \times PMLIQ$: $Low\ Quintile$	0.964***				
	(0.000)				
$Idio\_Volatility \times PMLIQ$: $High\ Quintile$	1.862***				
	(0.000)				
$Idio\_Volatility$		1.546***	1.745***	1.369***	1.609***
		(0.000)	(0.000)	(0.000)	(0.000)
$PMLIQ$		-0.319***	-0.222***	-0.220***	-0.279**
		(0.000)	(0.000)	(0.000)	(0.040)
$PMLIQ \times Idio\_Volatility$		0.094**	-0.190**	-0.207**	0.047
		(0.015)	(0.032)	(0.017)	(0.323)
$PMLIQ \times ETF\ Ownership_{lag}$					0.079
					(0.441)
$ETF\ Ownership_{lag}$			0.850***	0.819***	1.019***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times Idio\_Volatility$				-0.360***	-0.377***
				(0.000)	(0.000)
$ln(mktcap)$	7.595***	7.630***	7.383***	7.434***	8.446***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.202***	1.041***	1.275***	1.262***	1.091***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.478***	-0.509***	-0.382***	-0.380***	-0.492***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Market\ Beta$	0.452	0.134	0.539*	0.542*	0.124
	(0.135)	(0.628)	(0.079)	(0.077)	(0.650)
$ln(Amihud\ Illiquidity)\_lag$	7.381***	6.931***	7.566***	7.604***	8.181***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.963**	1.014**	0.852*	0.774*	0.903**
	(0.032)	(0.023)	(0.054)	(0.079)	(0.043)
N	672118	668974	668974	668974	663793
$adjusted\ R^2$	0.124	0.124	0.113	0.118	0.133

Panel B: Δ Idiosyncratic Volatility	DV = Stock Return				
	1	2	3	4	5
$PMLIQ$: $Low\ Quintile$	0.863***				
	(0.000)				
$PMLIQ$: $High\ Quintile$	-1.184***				
	(0.000)				
$\Delta Idio\_Volatility \times PMLIQ$: $Low\ Quintile$	1.990***				
	(0.000)				
$\Delta Idio\_Volatility \times PMLIQ$: $High\ Quintile$	3.500***				
	(0.000)				
$\Delta Idio\_Volatility$		2.424***	2.429***	2.068***	2.091***
		(0.000)	(0.000)	(0.000)	(0.000)
$PMLIQ$		-0.374***	-0.284***	-0.292***	-0.354***
		(0.000)	(0.000)	(0.000)	(0.009)
$PMLIQ \times \Delta Idio\_Volatility$		0.256**	0.267***	0.083	0.218**
		(0.015)	(0.009)	(0.348)	(0.049)
$PMLIQ \times ETF\ Ownership_{lag}$					0.100
					(0.353)
$ETF\ Ownership_{lag}$			0.738***	0.707***	0.709***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times \Delta Idio\_Volatility$				-0.818***	-0.820***
				(0.000)	(0.000)
$ln(mktcap)$	6.687***	6.179***	6.591***	6.538***	6.574***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.303***	1.293***	1.353***	1.334***	1.337***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.374***	-0.279***	-0.252***	-0.244***	-0.261***
	(0.000)	(0.002)	(0.005)	(0.005)	(0.003)
$Market\ Beta$	0.612**	0.579**	0.594**	0.590**	0.598**
	(0.044)	(0.043)	(0.036)	(0.037)	(0.035)
$ln(Amihud\ Illiquidity)\_lag$	6.679***	6.087***	6.870***	6.803***	6.821***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.982**	0.989**	0.961**	0.925**	0.927**
	(0.028)	(0.026)	(0.031)	(0.036)	(0.038)
N	669911	666787	666787	666787	661812
$adjusted\ R^2$	0.151	0.161	0.165	0.172	0.173

Table XI: Stock Turnover Volatility

Table reports time-series averages of coefficients from Fama-MacBeth (1973) cross-sectional regressions estimated each month over a sample period 2002 to 2019.

$$ret_i = \alpha + \sum_{q=Low}^{High} \beta_q \cdot PMLIQ_{i,q} + \gamma_1 \sigma_i^{turn} + \gamma_2 \sum_{q=Low}^{High} \left(\sigma_i^{turn} \cdot PMLIQ_{i,q}\right) + \sum_k \delta_k X_{ik} + \varepsilon_i$$

Dependent variable ret_i are monthly stock returns expressed in percent. $PMLIQ\_i,q$ are indicator variables that equal 1 for stocks that belong either to the lowest or highest quintiles, and 0 otherwise. Standard deviation of stock turnover, σ_i^{turn}, is computed each month from daily trading volume scaled by shares outstanding. Control variables X_{ik} defined in Table I are demeaned and standardized each month. Newey-West standard error of the time-series average of the difference between high and low quintiles are computed using a 12-month lag. p-values are reported in parentheses. ***, ** and * denote significance at the 0.01, 0.05 and 0.1 levels.

Panel A: Std Dev of Turnover	DV = Stock Return				
	1	2	3	4	5
$PMLIQ: Low\ Quintile$	0.911***				
	(0.000)				
$PMLIQ: High\ Quintile$	-1.206***				
	(0.000)				
$Std\_Turnover \times PMLIQ: Low\ Quintile$	0.716***				
	(0.000)				
$Std\_Turnover \times PMLIQ: High\ Quintile$	1.310***				
	(0.000)				
$Std\_Turnover$		1.014***	1.148***	1.074***	1.088***
		(0.000)	(0.000)	(0.000)	(0.000)
$PMLIQ$		-0.223***	-0.124*	-0.147**	-0.253*
		(0.000)	(0.052)	(0.029)	(0.091)
$PMLIQ \times Std\_Turnover$		0.164*	0.191**	0.082	0.142
		(0.050)	(0.032)	(0.377)	(0.154)
$PMLIQ \times ETF\ Ownership_{lag}$					0.083
					(0.396)
$ETF\ Ownership_{lag}$			0.925***	0.931***	0.932***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times Std\_Turnover$				-0.246***	-0.240***
				(0.000)	(0.000)
$ln(mktcap)$	7.498***	7.487***	8.150***	8.175***	8.208***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.282***	1.191***	1.254***	1.256***	1.260***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.433***	-0.384***	-0.350***	-0.344***	-0.363***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Market\ Beta$	0.582*	0.491	0.491	0.496	0.506
	(0.066)	(0.117)	(0.112)	(0.108)	(0.102)
$ln(Amihud\ Illiquidity)\_lag$	7.510***	7.392***	8.518***	8.547***	8.562***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	0.995**	0.995**	0.969**	0.961**	0.971**
	(0.027)	(0.025)	(0.029)	(0.032)	(0.032)
N	674089	670924	670924	670924	665374
$adjusted\ R^2$	0.120	0.119	0.123	0.125	0.126

Panel B: Ln Std Dev of Turnover	DV = Stock Return				
	1	2	3	4	5
$PMLIQ: Low\ Quintile$	0.862***				
	(0.000)				
$PMLIQ: High\ Quintile$	-0.977***				
	(0.000)				
$ln(Std\_Turnover) \times PMLIQ: Low\ Quintile$	1.018***				
	(0.000)				
$ln(Std\_Turnover) \times PMLIQ: High\ Quintile$	1.742***				
	(0.000)				
$ln(Std\_Turnover)$		1.664***	1.765***	1.689***	1.696***
		(0.000)	(0.000)	(0.000)	(0.000)
$PMLIQ$		-0.235***	-0.119**	-0.130**	-0.288**
		(0.000)	(0.025)	(0.017)	(0.043)
$PMLIQ \times\ ln(Std\_Turnover)$		0.099***	0.128***	0.024	0.081*
		(0.002)	(0.000)	(0.502)	(0.083)
$PMLIQ \times ETF\ Ownership_{lag}$					0.009
					(0.927)
$ETF\ Ownership_{lag}$			0.914***	0.967***	0.959***
			(0.000)	(0.000)	(0.000)
$ETF\ Ownership_{lag} \times ln(Std\_Turnover)$				-0.462***	-0.457***
				(0.000)	(0.000)
$ln(mktcap)$	7.983***	8.851***	9.512***	9.717***	9.744***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$ln(book/market)$	1.251***	1.096***	1.160***	1.151***	1.155***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.418***	-0.353***	-0.320***	-0.308***	-0.326***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
$Market\ Beta$	0.527*	0.292	0.292	0.293	0.305
	(0.092)	(0.337)	(0.331)	(0.329)	(0.307)
$ln(Amihud\ Illiquidity)\_lag$	8.054***	8.879***	10.006***	10.246***	10.250***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Constant$	1.014**	1.008**	0.982**	1.016**	1.017**
	(0.024)	(0.023)	(0.027)	(0.023)	(0.025)
N	674089	670924	670924	670924	665374
$adjusted\ R^2$	0.121	0.123	0.123	0.123	0.130

Appendix Table AI: Variable Definitions

$PMLIQ$: is the liquidity provided by primary market activity of APs is computed as $\hat{c}_{it}^{-1}$ where

$$\hat{c}_{it} = \underbrace{own_{it-1}}_{\substack{fund \\ ownership}} \cdot \underbrace{\textstyle\sum_j share_{ijt-1}^2}_{\substack{ownership \\ concentration}} \cdot \underbrace{\sqrt{\textstyle\sum_j \left[\omega_{ijt-1}\left(f_{jt}/a_{jt-1}\right)^2\right]}}_{\substack{volatility\ of \\ fund\ flow}}$$

own_{it-1} is aggregate fund ownership as a percent of the market capitalization of the stock, $share_{ijt-1}^2$ is a fund's ownership as a percent of aggregate fund ownership, and $\omega_{ijt-1} = share_{ijt-1}^2/\sum_i share_{ijt-1}^2$.

$ETF\ ownership$: is the percentage of total shares outstanding of stock i owned by all ETFs j at the end of month t. The number of shares of stock i owned by ETF is summed across ETFs j at the end of month t.

$$ETF\ Ownership_{i,t} = \frac{\sum_j shares\ outstanding\ owned\ by\ ETF_{i,j,t}}{total\ shares\ outstanding_{i,t}}$$

$\Delta ETF\ ownership$: month-by-month change in ETF ownership for a given stock.

$Market\ Capitalization$: is total shares outstanding multiplied by closing price at the end of the month.

$Book/Market$: book equity to shareholders' equity calculated following Daniel and Titman (2006).

$Momentum$: is the cumulative monthly return in the preceding 12-month period ending in one month prior to start of the month.

$Market\ Beta$: the coefficient on market return premium estimated with CAPM model using rolling 36 months returns.

$Amihud\ Illiquidity$: is daily absolute daily return divided by daily trading volume in millions of dollars averaged across days in the month.

$Turnover$: is French (2008) adjusted CRSP trading volume divided by shares outstanding.

$Asset\ Volatility$: is the volatility of return on the latent underlying assets of the firm estimated following Bharath and Shumway (2008).

$Stock\ Return\ Volatility$: computed following Jiang, Lee and Zhang (2005) as the standard deviation of daily returns of past 25 trading days.

$Idio\_Stock\ Return\ Volatility$: is the standard deviation of residuals estimated from a CAPM model using rolling 36- month window.

$Stdev\ Turnover$: the standard deviation of daily trading volume scaled by shares outstanding.

An Historical Loss Approach to Community Bank Stress Testing

Cao Fang, Timothy Yeager[1]

Abstract

We develop a top-down macro stress test that assesses a community bank's ability to withstand a severe and prolonged period of high credit losses. The model groups banks by geography and subjects them to the 90th percentile chargeoff rates that banks experienced between 2008 and 2012. Our historical loss approach better reflects patterns of community bank stress than econometric approaches that estimate the relationship between macroeconomic conditions and bank performance. We put all U.S. community banks at year-end 2017 through the test and highlight two results. First, banks are much better prepared to withstand an adverse shock than they were on the verge of the financial crisis because banks have shifted away from the riskiest loan types. Second, the Tax Cuts and Jobs Act of 2017 has increased bank insolvency risk from an adverse shock in 2018 because the higher bank capital is more than offset by the weaker automatic stabilizer effect from operating losses.

Keywords: Community banks, Stress testing, Financial Crisis, Loan diversification, Insolvency risk

JEL Codes: G01, G21

[1] tyeager@walton.uark.edu, 479-575-3944, Sam M. Walton College of Business, University of Arkansas, WCOB 475, Fayetteville, AR

I. Introduction

A macro bank stress test dynamically assesses a bank's insolvency risk and capital adequacy given an abrupt change in economic and financial conditions. Since 2009, the Federal Reserve has greatly expanded the importance of stress testing at the largest banking organizations. Annual results from the Dodd-Frank Act Stress Test (DFAST) and Comprehensive Capital Analysis and Review (CCAR) have effectively become the binding minimum capital requirements on large banking organizations, more onerous than the Basel III Capital Accord (Covas, 2017). Presently, community banks (banks with less than $10 billion in assets) are not required to conduct enterprise-wide stress tests required of larger organizations. However, each banking organization, regardless of size, is expected to analyze the potential impact of adverse outcomes on its financial condition (Board of Governors, 2012). Community banks, for example, are expected to stress test exposure to commercial real estate (CRE) lending (Board of Governors, 2006).

The primary objective of this paper is to introduce an historical-loss macro stress-testing model that assesses a community bank's ability to withstand a severely adverse yet plausible shock over a five-year horizon.[2] Although our historical-loss model differs from the more common econometric approach, it is more accurate in projecting patterns of distress at community banks similar to what they experienced in the years 2008-2012. The increased accuracy results from bypassing statistical approaches that introduce model error. Most stress-test models use econometric analysis to map historical changes in macroeconomic variables onto bank performance; however, researchers have shown that the predictive content of macroeconomic variables in forecasting large bank performance is weak, introducing a wide confidence band around point estimates (Guerreri and Welch, 2012;

[2] The stress test generates reports for every U.S. community bank and thrift and is freely available at [removed]. The model is run in Microsoft Excel and updated annually.

Grover and McCracken, 2014). This problem is exacerbated for community banks because they operate in state and local markets where economic data are relatively poor (Barth et al., 2018).

Our historical-loss stress test bypasses the econometric mapping process by exposing each community bank to the 90th percentile chargeoff rates experienced by banks in the local geographic market of its headquarters in the years 2008-2012, a period encompassing the financial crisis and Great Recession.[3] This approach directly links each bank's projected stressed chargeoffs to its local market, and it naturally captures nonlinear outcomes that confound standard econometric procedures. The relative simplicity of our model is also a helpful feature for community banks because it is easy to use and interpret (Schmieder, Puhr, and Hasan, 2014). We assess the in-sample validity of the model by putting all U.S. community banks through the stress test based on their financial conditions at year-end 2007. Three-fourths of the community banks that failed between 2008 and 2012 also fail or experience dangerously low capital levels during the stress test.

A second objective of this paper is to assess the ability of present-day banks to weather a severely adverse shock. We run our stress tests on bank condition at year-end 2017 and find that banks are much better positioned for a severe downturn in 2017 than they were in 2007 because the riskiest banks are no longer in business, and construction and land development (CLD) loan concentrations are lower. In the extreme, substituting all CLD loans with nonfarm nonresidential (NFR) commercial real estate loans brings large diversification benefits; the number of community banks projected to fail declines by 68%. We also show that the Tax Cuts and Jobs Act of 2017 (TCJA) has diminished the ability of banks at year-end 2018 to survive an adverse shock because the higher capital is more than

---

[3] A variety of evidence shows that community banks lend and take deposits locally. The 2018 FDIC Small Business Lending Survey shows that 82.3 percent of small banks (less than $10 billion in assets) selected the city, county, or MSA as their relevant trade area. In addition, the FDIC Summary of Deposits data show that as late as 2018, banks with less than $500 million in assets held more than three-quarters of their deposits in their headquarters county. Similarly, urban banks of the same size held 92 percent of their deposits in their headquarters MSA. Finally, Petersen and Rajan (2002) show that although the distance between bank lender and small business borrower increased between the early 1970s and 1990s, the median distance increased from just 2 miles to 5 miles. Agarwal and Hauswald (2010) analyze a confidential sample of small business loans from a large U.S. bank and find a median (mean) distance of 2.62 (9.9) miles for accepted loans.

offset by weakened automatic stabilizers from net operating losses. Among other things, the TCJA forces banks to recover tax benefits from operating losses through deferred tax assets (DTAs), which are excluded from Tier 1 capital.

The paper proceeds as follows. Section 2 discusses weaknesses of an econometric approach for community bank stress testing, and Section 3 explains our historical loss methodology. Section 4 assesses the performance of U.S. community banks after subjecting them to the stress tests in 2007, 2017, and 2018. Section 5 conducts in-sample model testing. Section 6 evaluates the potential diversification benefits from bank reallocation of loan portfolios, and Section 7 concludes.

II. Weaknesses of the Econometric Approach

A macro stress test projects the effects of an adverse macroeconomic scenario onto bank performance. Because credit risk is the central focus, the critical assumption is the projection of multi-period bank chargeoffs. Most researchers and practitioners use econometric techniques to estimate the historical relationship between macroeconomic variables and bank performance (Covas, 2014; Hirtle et al., 2016; Kapinos and Mitnik; 2016). The econometrician regresses bank net income components and net chargeoffs by loan type on economic variables such as real estate prices, GDP growth, and unemployment rates. The coefficients on the variables (or principal component indices) are then used to project the average changes in chargeoffs, net income, and capital for each bank given the hypothetical change in the economy. A related methodology is a vector autoregression (VAR) where bank performance is projected from an interdependent system of regressions on banking and economic variables (Hall et al., 2011; Jacobs, 2016). The key benefit is that a VAR captures predictable variations based on dynamic correlations. VARs, however, require relatively long time series to produce statistically reliable results, and the minimum sample size grows with the number of variables in the system.

Econometric approaches to macro stress testing pose two significant challenges. First, standard techniques maintain an assumption of conditional normality, making nonlinearities and tail events difficult to capture. Researchers have adopted different techniques for large bank models to address this challenge. Covas et al. (2014) use a quantile regression approach to capture nonlinearities. Jacobs (2016) shows that a Markov Switching VAR is better suited to capture extreme events than the standard VAR model, and Kapinos and Mitnik (2016) use an optimal grouping strategy where slope coefficients differ among dynamic groups of banks. These procedures, however, potentially introduce more model error and add complexity that must be weighed against a simpler approach necessary for community banks.

A second more serious challenge is that the empirical connection between macroeconomic data and bank performance is tenuous. Guerreri and Welch (2012) find for a sample of large BHCs that macroeconomic variables have little predictive power in forecasting banking variables; confidence bands around the forecasts are too wide to distinguish between adverse and severely adverse macroeconomic scenarios. Grover and McCracken (2014) investigate the usefulness of factor-based methods for assessing industry-wide bank stress. They find that none of their factor models forecasts net chargeoffs better than a random walk. The authors then use the factors to measure the degree of stress faced by the banking industry in each of the 2014 CCAR baseline, adverse, and severely adverse scenarios. Consistent with Guerreri and Welch (2012), they find little difference in projected bank outcomes between the adverse and severely adverse scenarios. In fact, net chargeoffs are often higher after one year under the adverse scenario than the severely adverse scenario.

The empirical relationship between macroeconomic data and net chargeoffs is even more tenuous for community banks. Because these banks are geographically concentrated, local and state economic data may be more relevant than national economic data, yet data availability is relatively sparse, and the quality relatively poor. Moreover, the connection between bank performance and local economic conditions is unclear. Yeager (2004) finds small differences in the performance of community banks in counties that suffered spikes in unemployment rates in the early 1990s relative to similar banks in counties that did not suffer such spikes. Barth et al. (2018) test the relative importance of 364 macro and banking variables in predicting chargeoffs for two large banks and two community banks. They find that bank predictors dominate all macro predictors in forecast accuracy, and state macro predictors slightly outperform national macro predictors, both for large banks and community banks.

DeYoung and Fairchild (2018) develop a community bank stress test using an econometric methodology based on the large-bank model of Hirtle et al. (2016), which estimates the relationship between bank performance and macro conditions by regressing bank performance metrics on bank-specific characteristics and national macroeconomic variables for the years 1991-2015. As robustness, they supplement regressions with principal components derived from state-level economic data. The contribution of state-level data to the R^2 of the regressions, however, is modest. The authors run stress tests separately on community banks with assets above and below $500 million, and they find the smaller community banks are essentially unaffected by the adverse scenarios. This result is troubling because 74% of community banks that failed from 2008-2012 had less than $500 million in assets, suggesting that the regression approach poorly captures the relationship between macroeconomic conditions and bank performance.

In addition to nonlinearities and model error, econometric approaches have difficulty replicating the cyclical chargeoff patterns of banks observed through a business cycle. For example, as the economy deteriorated in 2008, bank chargeoffs rose slowly, peaked in 2009 and 2010, and tapered off thereafter. In contrast, econometric approaches typically impose shocks that taper off immediately.

III. Historical Loss Stress Test Methodology

We adopt an historical-loss stress-test methodology that bypasses regression estimation and the ensuing model error, naturally incorporates tail events and local economic conditions, requires only five years of annual data, and intrinsically incorporates cyclical shocks to credit quality that build and taper through time. We group community banks into geographic markets and run a five-year simulation that subjects each bank to its market's 90th percentile net chargeoff rates by loan type for each year from 2008-2012. This five-year horizon is chosen because it captures the deterioration and recovery of bank balance sheets from the financial crisis and Great Recession.[4] A limitation of the historical loss approach is that, unlike CCAR and DFAST, it is not adaptable to changing adverse economic scenarios. However, the severe stress that banks experienced from 2008-2012 is a plausible and reasonable scenario benchmark even in today's healthier banking climate. Indeed, reducing the severity of the shock as banking conditions improve can introduce additional sources of procyclicality.[5] The loss rate in the model is easily customized, however, so users can apply lower loss rate percentiles to run less severe scenarios.

III.1. Geographic Banking Markets

Each community bank is assigned to one market based on the location of its headquarters. We define geographical market boundaries to ensure that a reasonably large number of banks exists in each market. Banks headquartered in an MSA (urban banks) with at least 30 banks headquartered in that MSA comprise their own market. Banks in MSAs with less than 30 banks are grouped by state to form a market if there are at least 20 such banks across the state. All rural banks (not headquartered in MSAs) in a state form a market if there are at least 20 rural banks. If there are fewer than 20 urban

[4] As robustness, we use loss rates from 1991-1995 and the results show, as expected, that the number of banks that undergo severe stress decline between 64% and 78%. Consequently, we do not view that time period representative of a severely adverse scenario.
[5] Some large banks subject to CCAR and DFAST commented to the Federal Reserve that the scenarios should be less severe to be more in line with historical post-war recessions. The Federal Reserve responded that, by design, the severity of the scenarios increases as economic conditions improve to limit sources of procyclicality in the stress tests. See "Amendments to Policy Statement on the Scenario Design Framework for Stress Testing," Federal Register, February 28, 2019: 84(40) 6654.

banks from the smaller MSAs in the state or fewer than 20 rural banks in the state, those urban and rural banks are combined into one market. In all, there are 46 unique markets for community banks headquartered in rural markets. For most years, 32 of those markets consist only of rural community banks headquartered in the same state, and 14 markets consist of rural and urban community banks in the same state. Similarly, there are 66 unique urban markets for community banks headquartered in urban markets. For most years, 16 markets consist of banks in the same MSA, 33 markets consist of urban banks in the same state, and 17 markets consist of rural and urban community banks in the same state.[6]

III.2. Loss Rates at the 90th Percentile

After establishing markets, we impose on each bank the 90th percentile net chargeoff rate for each loan type experienced by banks in its respective market for each of the five years 2008-2012. Table I lists by loan type mean 90th percentile net chargeoff rates experienced by rural and urban markets, respectively, for each year 2008-2012. Annualized chargeoff rates for each bank are computed quarterly, averaged by year, and winsorized at the top and bottom 1% to eliminate outliers. The 90th percentile chargeoff rates from each market are then averaged by year and MSA status. In general, urban markets experience far higher chargeoff rates than rural markets, and loss rates for 'other' CLD (CLD-OTH) loans are particularly high, peaking at 12% in 2010.

Given that there are 11 loan types in the model and actual chargeoff rates across loan types within a bank are not perfectly correlated, the joint probability is extremely low that a bank will experience the 90th percentile loss rate in every loan type, which at first glance suggests that we overstate the severity of the shock to the bank. Indeed, the median failed bank between 2008 and 2012 had four loan categories that exceeded the threshold prior to failure. However, it is more

[6] Six states have fewer than 20 community bank headquarters, so we ensured that the 90th percentile loss rates from these markets are consistent with the other markets. On average, the loss rates are lower in the markets with fewer than 20 banks.

important for the stress test to accurately represent the actual loan *portfolio* chargeoff rates experienced by banks during the crisis years rather than the chargeoff rates of each loan category. Even if a bank incurs its market's 90^{th} percentile chargeoff rate for each loan type, its chargeoff rate for the portfolio as a whole may be above or below its market's portfolio 90^{th} percentile chargeoff rate depending on the bank's loan composition relative to the market. The 90^{th} percentile chargeoff rates on some loan types (such as construction loans) between 2008 and 2012 were much higher than others, and a bank with a greater share of those loan types could experience a portfolio chargeoff rate above the market's 90^{th} percentile even if loss rates on other loans were well below the 90^{th} percentile.[7]

We show that our simulated portfolio loss rates are reasonable relative to the actual chargeoff rates experienced by community banks. Because we observe the actual distributions of portfolio chargeoff rates in each market and year between 2008 and 2012, we can compute the projected chargeoff rate percentile for each bank from the stress test results. As we show later, in-sample stress test results produce portfolio chargeoff rates that on average range between the 86^{th} and 90^{th} percentiles of actual portfolio chargeoff rates. In addition, imposing the same 90th percentile loss rate on all loan types alleviates the concern that we are choosing loan types ex-ante that will perform better than others.

Although survivorship bias is present in the computation of the 90^{th} percentile loss rates, it is appropriate for our purposes to retain this bias. Survivorship bias arises because all failed banks are excluded from the loss rate calculations in the years after they fail, and because banks that failed during the 1st quarter of a given year are excluded from the loss rate calculations within that same year. Loss rates from banks that failed after the 1st quarter of the year are included in that year's loss rate distribution because we compute each bank's annualized net chargeoffs each quarter from the Call

[7] As robustness, we tested the model with stress tests ranging from the 85^{th} to 95^{th} percentiles, and we found that no sharp discontinuities existed around the choice of the 90^{th} percentile.

Reports and average them over the calendar year. The purpose of selecting the five-year loss rate window from 2008 to 2012 is to capture the deterioration and recovery of banks over a full economic cycle. Including the loss rates of previously failed banks each year alters the loss-rate distributions and amplifies the persistence of the simulated shock, weakening the natural recovery process.[8]

III.3. Model Dynamics

For the stress test, the initial condition of each bank is taken from its annualized year-to-date Call Report data as of the fourth quarter of the year being tested. Inputs include loan amounts, average loan yields, loss rates, and other information obtained from publicly available Call Reports. The bank-specific simulation input worksheet for the fictitious "Sample Community Bank" is presented in Appendix A. The simulation then projects financial ratios five years forward after applying the relevant chargeoff rates.

Assets in year *t* consist of securities, federal funds sold, interest-bearing balances, loans (L), and loan loss reserves (LLR).[9] All liabilities are represented as deposits (D), and shareholders' equity (E) is the difference between assets and liabilities. We lump federal funds and interest-bearing balances with securities (S) so that the balance sheet is represented as:

$$S_t + L_t - LLR_t = D_t + E_t \tag{1}$$

We assume that banks reinvest all principal and interest payments in the same asset categories. Consequently, securities grow according to:

$$S_{t+1} = S_t(1 + a) \tag{2}$$

[8] As robustness, we ran simulations that included loss rates from banks that failed in the first quarter of the year by using their final reported loss rates from the fourth quarter of the previous year. Projected bank failures increased by small amounts, easing concerns that same-year survivorship bias has large effects on the results.
[9] Non-earning assets are excluded for ease of exposition. Because the core simulation model is similar to that in Hall et al. (2011), we draw heavily from that approach.

where α is the target growth rate of assets, which we set at 3%. Charged-off loans, however, are not reinvested so that loans (and hence, total assets) decrease by the amount of chargeoffs.[10] The bank's *j* loan categories in its portfolio grow through time as:

$$L_{t+1} = \sum_j (1 - c_{j,t+1}) L_{jt}(1 + a) \tag{3}$$

where $\boldsymbol{c_j}$ is the annual charge-off rate for loan category *j*. Because the stress test focuses exclusively on credit risk, we do not explicitly change interest rates over the simulation horizon. To the extent that interest rates affected chargeoffs over the 2008-2012 period, some of their dynamics are captured in the historical loss rates.

Banks use provision expense (*P*) to offset exactly net chargeoffs (*LS*) in the current year if net chargeoffs are positive, but provision expense is zero if net chargeoffs are negative. We also cap the loan loss reserve to total loan ratio at 1.5% to allow for banks to draw down high initial loan loss reserves before replenishing them with new provisions.[11] In addition, banks add to provisions an amount equal to the realized loan growth rate as:

$$P_t = \max[0, LS_t] + LLR_{t-1} \cdot (L_t/L_{t-1} - 1) \tag{4}$$

Loan loss reserves, then, change through time according to:

$$LLR_t = LLR_{t-1} + P_t - LS_t \tag{5}$$

Net income is computed each year as:

$$NI_t = r_s S_t + \sum_j r_j L_{jt} - r_d D_t - NNE_t - P_t - T_t \tag{6}$$

[10] Sensitivity analysis of the target growth rate shows that it has large effects on model outcomes. Higher growth rates induce more projected bank failures because more assets reduce the capital ratio. We choose a 3% growth rate because it is similar to the long-run U.S. nominal GDP growth rate so that bank assets and economic growth have similar long-run trends. The actual loan growth rate should fall during a period of stress given the economic decline. Interestingly, the actual community bank growth rate (winsorized at the top and bottom 10%) averaged 3.0% annually between 2008 and 2012, while the 2017 simulated loan growth averaged 1.3%.

[11] Call Report data show that at year-end 2007, the 75th percentile of the ratio of LLR to total loans was 1.5%. Consequently, we assume that most banks would be comfortable drawing down their reserves to that level, but they would add to provisions to rebuild their reserves below that threshold. This assumption has little effect as it prevents just a handful of banks from dropping below critical capital thresholds.

where $\boldsymbol{r_s}$ is the average rate on securities, $\boldsymbol{r_j}$ is the rate on loan *j*, $\boldsymbol{r_D}$ is the average rate on deposits, *NNE* is net noninterest expense (noninterest expense less noninterest income), and *T* represents taxes. Deposit interest expense, noninterest expense and noninterest income are assumed equal to their initial percentages of total assets and they change in proportion to the bank's total assets. Taxes are assumed to be 35 percent of operating income, though we also run the tests using the 2018 corporate tax rates of 21%.

Finally, the dividend payout ratio (*d*) is assumed equal to the initial ratio of dividends to net income (NI); however, dividend payments are set to zero if net income turns negative so that

$$DIV_t = \max\,[0, d \cdot NI_t] \tag{6}$$

Retained earnings (RE) equal net income less dividends, and they boost equity (E) such that

$$E_t = E_{t-1} + RE_t \tag{7}$$

Deposits are assumed to automatically adjust each period to balance the balance sheet, as in Equation (1). Figure I provides a flow chart that summarizes the simulation logic.

III.4. Capital Thresholds

The most important metric from the stress test results is a bank's Tier 1 Leverage (T1Lev) ratio, or Tier 1 capital divided by Tier 1 average assets. Because the simulation computes equity directly, the T1Lev ratio for each bank each year is derived by subtracting the initial (Y0) Tier 1 capital from equity, and initial (Y0) Tier 1 average assets from total assets, and we hold those differences constant throughout the simulation. For a given simulation, we track the number of banks where the T1Lev ratio falls below 2% and 6%, respectively. The 2% threshold mimics the Prompt Corrective Action (PCA) guidelines that define a bank as "critically undercapitalized" if its tangible equity is equal to or less than 2% of total assets.[12] If the capital deficiency is not corrected, a critically undercapitalized bank must be placed into receivership within 90 days by regulators.

We also track the number of banks where the T1Lev ratio falls below 6% to identify banks falling into a dangerous capital zone that signaled high insolvency risk during the crisis and Great Recession. Although the 6% threshold is above the 5% ratio that PCA defines as "well capitalized," it is abundantly clear that the 5% ratio was too low to flag banks with high insolvency risk. Cole and White (2017) show that the mean equity ratio of banks closed between 2007 and 2014 was 6.4% one year prior to failure, and the mean equity ratio did not fall below 2% until one quarter prior to failure. In addition, the Government Accounting Office (2011) concluded that the PCA framework did not prevent widespread losses to the deposit insurance fund. Consequently, the 6% T1Lev threshold represents a reasonable lower bound ratio signaling insufficient capital even though it exceeds PCA guidelines.

IV. U.S. Community Bank Stress Test Results

IV.1. Stress Tests Results for 2017

We run stress tests on all 4,846 community banks at year-end 2017 and simulate results for the years 2018-2022. Aggregate results are presented in Table II. (The simulation output for a representative bank is shown in Appendix B.) Loan loss rates for a given bank come from the 90th percentile chargeoff rates experienced by banks in the same market between 2008 and 2012, and the initial condition of the banks in Year 0 is taken from financial data at year-end 2017.

The top panel of Table II shows that the mean bank begins the simulation well capitalized with a T1Lev ratio of 11.5% and a median ratio of 10.4%. Despite the severe shocks that hit the banks, mean capital ratios remain high over the five-year horizon; the mean ratio in Year 5 (2022) is 10.2%.

[12] "Tangible equity" is defined by the regulators as Tier 1 capital plus outstanding cumulative perpetual preferred stock (including related surplus) not already included in Tier 1 capital. Changes to the Call Report after 2014 do not allow us to precisely measure the cumulative perpetual stock not already in Tier 1 capital, so we use the T1Lev ratio as a proxy for tangible equity.

Just 153 of the 4,846 banks (3.2%) T1Lev ratios that fall below 2% during the forecast horizon, implying that they would be closed by regulators in the absence of additional capital injections. Another 563 banks (11.6%) have T1Lev ratios that fall below 6%. Not surprisingly, the stress test results show that bank profitability plummets. The bottom panel of Table II lists mean ROA, which reaches its nadir in Year 3 (2020) at -9bp before improving in Years 4 and 5. In Y3, 2,246 banks (46.4%) have negative earnings.

The middle panel of Table II reports mean net chargeoff rates from the 2017 stress tests, and they peak in Year 3 at 2.3%. To assess the plausibility of these loss rates, we examine how closely the projected portfolio chargeoff rates are correlated with the actual 90th percentile portfolio chargeoffs community banks experienced between 2008 and 2012. Mean chargeoff rates from the 2017 stress tests are plotted as dark-shaded columns (ST2017) in Figure II. The light-shaded columns (P90) display means of the actual 90th percentile portfolio chargeoff rates for all U.S. community bank markets between 2008 and 2012. The mean chargeoff rates from the 2017 stress tests are consistently lower than actual 90th percentile chargeoff rates. Figure II also plots as a dashed line (ST2017PCTL) the percentiles of mean chargeoff rates from the 2017 stress tests relative to actual chargeoff rate distributions between 2008 and 2012. The right-hand axis represents percentile ranking and shows that mean projected chargeoff rates lie between the 86th and 90th percentiles. In sum, applying the 90th percentile chargeoff rate to each loan category produces simulated loan portfolio chargeoffs at reasonably severe levels.

As a further check, we compare our mean loss rates from 2017 with the mean loss rates from the 2017 DFAST severely adverse scenario taken from Table VII of the report (Board of Governors, 2017). Given the vast differences in the types of banks and the simulation processes between the two stress tests, our sole objective is to observe whether the loss rates are reasonably similar. To make the comparisons more appropriate, we report loss rates from the 2017 community bank simulation using only urban banks because DFAST banks operate in urban areas. In addition, DFAST loss rates are computed as cumulative losses over the nine-quarter horizon divided by the average loan balances over the same period. In contrast, our loss rates are computed annually, so we recompute them as Y1+Y2+Y3/4. In other words, we sum community bank loss rates from 2008, 2009, and one-fourth of 2010. Table III lists the nine-quarter mean loss rates for the 2017 stress tests from both models. DFAST loss rates are 2.4 percentage points higher for CRE loans, but the community bank simulation shows higher loss rates for consumer, C&I, and Agriculture loans.[13] Overall nine-quarter portfolio loss rates are 5.0% for the urban community bank simulation compared with 5.8% for DFAST. We conclude that although loss rates across the loan categories are somewhat different between the models, the overall portfolio chargeoff rates are similar, providing further evidence that the community bank simulation results are reasonable.

IV.2. Comparison of 2007 and 2017 Stress Test Results

The 2017 stress tests show that most community banks at year-end 2017 can weather a severe downturn and sustain high capital ratios. It is interesting to ask how community banks in 2007 would have fared the stress test.[14] Table IV displays stress test results for the 7,125 community banks at year-end 2007, and performance is much worse than the 2017 results. The number of projected failed banks in Y5 is 762, or 10.7% of all community banks relative to 3.2% in 2017. In addition, the percentage of banks with T1Lev ratios projected to fall below 6% is 23.0%, double the 11.6% value in 2017.

[13] We use the "other consumer" category from DFAST to exclude credit cards because community banks extend few credit card loans.

[14] Because many banks did not report the subdivided NFR and CLD components separately in the 2007 Call Reports, we estimate the component values of loans and chargeoffs at year-end 2007 by applying the percentages from March 2008. See Appendix C for details.

We identify two explanations for the more favorable stress test outcomes in 2017 than in 2007.[15] First, many of the riskiest banks in 2007 dropped out of the sample by 2017.[16] Nearly two-thirds of the 762 banks projected to fail the 2007 stress tests were no longer in business in 2017, and these banks had loan portfolios with high default risk. Table V summarizes several stress test scenarios by listing the percentage of banks that dropped below the 2% (and 6%) T1Lev threshold. Row 1 shows that in the 2007 baseline stress test, 10.7% of banks dropped below the 2% threshold. Rows 2 and 3, however, show that the number jumps to 19.5% for banks in 2007 that did not exist in 2017, and it falls to 4.9% for banks that existed in both years. This disparity exists because the 2,509 banks that existed in 2007 but not in 2017 had higher inherent credit risk than the 4,616 banks that existed in both years. Figure III plots the mean differences in 2007 loan shares by bank status in 2017. Banks that did not exist in 2017 held nearly 7 percentage points more in CLD-OTH and CLD-RES loans, and 3.7 percentage points more in NFR-OTH and NFR-OWN loans than banks still in the sample in 2017.

Removing banks that did not exist in 2017 from the 2007 bank stress test accounts for most but not all the difference in insolvency risk between 2007 and 2017. The projected failure rate from the 2007 stress test that includes only banks that also existed in 2017 is 4.9% (Row 3 of Table V), still higher than the 3.2% (Row 4) in the baseline 2017 simulation. A second explanation for the improved stress-test performance in 2017 relative to 2007 is that the banks that existed in both years adjusted loan portfolios during the interim 10-year period away from loan types with high default rates such as CLD. Figure IV shows this shift. Panel A plots major loan category shares for all banks that existed in both years, but it shows modest changes between 2007 and 2017. The CRE share, for example, increased from 41% to 45% of loans while consumer and commercial and industrial loan shares

[15] One potential explanation is that community banks held more securities and fewer loans in 2017 than 2007. Distributions of loan-to-asset ratios, however, are similar across those years.
[16] Of the 2,509 banks that disappeared from the sample, 18% failed, 2% converted to thrifts or exceeded the $10 billion asset threshold for a community bank, and the remaining 80% were acquired or converted to a bank branch.

declined by 3% and 2%, respectively. Panel B of Figure IV, however, shows larger changes within the CRE portfolio. CLD loans declined by 11% from 2007 to 2017, NFR loans increased by 4%, farm loans increased by 4%, and multifamily loans increased by 3%.

To estimate the effects on stress test outcomes from loan portfolio shifts between 2007 and 2017, we first run stress tests on the 4,616 U.S. community banks in 2017 that also existed in 2007. Row 5 of Table V shows that 2.5%, or 117 banks are projected to fail. We then adjust the loan shares of each bank in 2017 to equal its loan shares in 2007. Row 6 of Table V shows a sharp rise to 4.5%, or 206 projected failures, which explains most of the 4.9% (Row 3) projected failure rate of banks in the 2007 simulation. We conclude that the reduction in CLD loan shares has greatly reduced bank insolvency risk over the last decade.

Taken together, these results show that banks had much lower insolvency risk in 2017 than in 2007 because most of the riskiest banks in 2007 no longer existed in 2017, and because the existing banks in 2017 shifted away over the last decade from loan categories with relatively high default risk.

IV. 3. Effects from the New Tax Law

We also use the stress test to examine effects on community bank insolvency risk from the Tax Cuts and Jobs Act (TCJA) signed into law in December 2017. The new law lowers the corporate tax rate, eliminates net operating loss (NOL) carrybacks, and limits NOL deductions to 80% of taxable income. Each of these changes weakens the automatic stabilizers that banks with negative NOLs received under the previous tax law. We show that TCJA has increased community bank insolvency risk in 2018 relative to 2017 because, although banks boosted capital ratios in response to the tax cuts, the higher capital was insufficient to offset the weakened automatic stabilizers.

The TCJA weakens automatic stabilizers that help offset bank losses in adverse conditions. Imagine a bank that has net operating income during the years Y1 through Y3 of -\$100, -\$100, and +\$100. Relative to the previous tax law, automatic stabilizers are weakened in three ways. First, the new law lowers the corporate tax rate to 21% from 35%, which reduces the NOL tax benefit in Y1 and Y2 to $21 ($100 x 0.21) from $35 ($100 x 0.35). Second, TCJA eliminates NOL carrybacks so banks must exclusively use carryforwards, which has implications for cash flow and regulatory capital. Under the previous tax rules, a bank with a $100 NOL in Y1 and again in Y2 could have applied carrybacks for taxes paid in the previous two years to receive a cash-based tax benefit of $35 each year, which reduces annual after-tax loss to $65.[17] Under the new rules, a bank would receive an accrual-based tax benefit each year of $21 called a deferred tax assets (DTA), which represents a claim on reduced cash-based tax payments in future profitable years. Importantly, the Basel III Capital Accord excludes those DTAs from Tier 1 capital because severely distressed banks may be unable to generate future positive operating income to utilize the DTAs. Although the DTA asset increases equity in the year of the NOL, regulatory capital ratios are unaffected until the DTAs are utilized. Third, TCJA limits carryforwards to 80% of taxable income. When the bank in our example earns net operating income of +$100 in Y3, it can utilize the accumulated DTAs ($42 over Y1 and Y2) to offset the taxable income. Previously, a bank using carryforwards could utilize $35 in DTAs to offset the full Y3 tax obligation. Under TCJA, the bank can offset only 80% of the taxable income, or $16.80 (80% x $100 x 21%) with DTAs, leaving a tax obligation of $4.20.

We compare the effects from the 2017 stress test on T1Lev ratios under the current and previous tax laws. Under the new tax law, DTAs that result from NOLs during the simulation are subtracted from the T1Lev ratio. Row 7 of Table V shows the sharp increase in insolvency risk from the new tax law relative to the 2017 baseline (Row 4). The percentage of banks with T1Lev ratios below 2% more than doubles from 3.2% to 6.8%. The increased projected bank insolvency arises for two reasons. First, as discussed above, the lower tax benefit from operating losses combined with the

[17] Although the previous tax law limited NOL carrybacks to two years, in November 2009 Congress extended carrybacks to five years for most banks for losses in 2008 and 2009. Consequently, banks could offset losses in 2008 or 2009 with tax payments made from 2003 to 2007.

exclusion of DTAs from Tier 1 capital reduce regulatory capital ratios. The effect on regulatory capital can be seen in Row 7 of Table V as just 4.5% of banks cross the 2% equity-to-asset threshold compared with 6.8% for the T1Lev ratio. Second, the stress test run in 2017 affords banks no time after the change in the tax law to accumulate more capital even though net income should increase due to reduced tax payments. The fixed dividend payout ratio built into the model may also slow capital build-up because banks are assumed to keep payout ratios constant for all years with positive earnings rather than decreasing them as net income rises. But even if payout ratios remain unchanged, higher industry operating income would increase capital over time. It is an empirical question as to whether banks will accumulate more capital in the low-tax environment to offset the weaker automatic stabilizers.

To answer this question for the year 2018, we first examine the effects of TCJA on banks' capital in 2018 relative to the prior two years. Figure V plots ROA, the dividend payout ratio, and the T1Lev ratios for community banks between the 10^{th} and 90^{th} percentiles. We average the values for banks in 2016 and 2017 to reduce the likelihood that banks already began to adjust their behavior in anticipation of the tax cuts (Wagner, Zeckhauser and Ziegler, 2018). Panel A shows, as expected, that bank earnings increased consistently in 2018 along the distribution. Panel B shows that dividend payout ratios are slightly higher in 2018 for banks with payout ratios less than 20%, but they decline for banks with payout ratios above 20%. The net result is that higher income combined with lower dividend payout ratios for most banks boosted the median Tier 1 leverage ratio in 2018 by 17bp, which we observe in Panel C of Figure V.

Even given that community banks responded to the TCJA by increasing capital ratios in 2018, we examine whether the increase is sufficient to offset the weakening automatic stabilizers. We run stress tests in 2018 that include only banks that existed both in 2017 and 2018. Row 8 of Table V presents the stress test results of the 4,598 community, and it projects that 5.9% of banks fall below the 2% T1Lev ratio threshold. The projected failures are below the 6.8% of banks projected to fail in the 2017 simulation under TCJA reported in Row 7, which indicates that the higher T1Lev ratios have reduced insolvency risk relative to what it would have been without an increase in capital. The relevant comparison, however, is with the much lower 3.2% of banks that failed the baseline 2017 stress test reported in Row 4. At year-end 2018, TCJA has increased community bank insolvency risk from a severely adverse scenario because the increased capital ratios are not sufficient to offset the weakened automatic stabilizers.

V. In-Sample Model Performance

The value added from a stress test is the ability to identify banks that are the most vulnerable to a sudden adverse shock. Out-of-sample testing of our model is not yet possible because the parameters are based on recent experience and, fortunately, banks have not experienced another shock like the 2007-2009 financial crisis. However, we can conduct in-sample tests by comparing stress-test projections with the actual performance of community banks at year-end 2007 that were not acquired in the years 2008-2012. We expect the stress tests to detect most banks that failed during the period (low Type I error). However, because we apply the 90th percentile loss rates to all banks, the model should identify more banks as distressed than those that actually became distressed (high Type II error).

The stress test identifies a failed bank as one that has a T1Lev ratio below 2% at some point during the simulation. A Type 1 error occurs when the stress test fails to identify a failed bank. Of the 386 community banks from year-end 2007 that failed between 2008 and 2012, the stress test identifies 215 of those banks, resulting in a Type 1 error of 171 banks (45%). Book capital ratios were commonly overstated at many banks that failed during this period because banks were reluctant to recognize losses in a timely manner (Garcia, 2010; GAO, 2011). If we assume that banks with simulated T1Lev ratios less than 6% also were likely to fail, the stress test flags an additional 72 failed banks, reducing the Type I error to 99 banks (26%). For these remaining Type I banks, the average minimum simulated T1Lev ratio is 10.4% and the average minimum ROA is -1.1%.

A Type II error occurs when the model incorrectly identifies a healthy bank as distressed. The 2007 stress tests forecasts a total of 762 banks with T1Lev ratios below 2%. Of those banks, 215 failed and 160 were closed or acquired between 2008 and 2012, resulting in 387 banks projected to fail that did not fail. Among those, 80 banks were enrolled in the TARP program, which may have been used to recapitalize the banks to avoid failure. Excluding those 80 banks, the Type II error is 307 banks (40%). For Type II banks, the average minimum simulated T1Lev ratio is -2.1% and the average minimum ROA is -3.7%.

Our 2007 stress test results accurately project failure risk by community bank size. Recall that 74% of community banks that failed from 2008-2012 had less than $500 million in assets. Fully 79% of banks that our model projects to fail had assets in 2007 less than $500 million.

Another approach to measuring in-sample performance of our stress test is to compare the model's results with traditional early warning signals of bank distress. Stress tests differ from early warning signals because they subject all banks to adverse shocks while early warning signals are static indicators designed to detect banks with high default risk at a point in time. Nevertheless, it is reasonable to expect overlap between the banks flagged by early warning signals and those that perform poorly in stress tests.

A simple and potentially powerful early warning signal is the T1Lev ratio. Banks with higher capital cushions, all else equal, can absorb more losses before failure. How likely is the stress test to project banks with the lowest leverage ratios in Year 0 to fail? Using year-end 2007 Call Report data as Year 0, Table IV shows that the spearman rank correlation coefficient between actual Y0 T1Lev ratios and projected Y5 T1Lev ratios is 0.64. The same correlation for the 2017 simulation is 0.75.

In both years, a strong correlation exists between the initial Tier1 leverage ratio and the projected capital ratio in Year 5.

A more robust early warning signal is the Federal Reserve's SEER failure probability model, designed to predict the likelihood of bank failure over the subsequent two years (Cole, Cornyn, and Gunther, 1995). Each bank's failure probability is derived from a multinomial probit regression of bank failures in the mid-1980s through the early 1990s. The coefficients from this model are confidential, but Miller et al. (2015) replicate the model and show that the so-called dated failure probability (DFP) signal was the most accurate of a host of early warning signals for detecting bank failures from 2009 through 2012.[18] We rank banks by failure probability from highest to lowest (riskier banks have lower ranks) and compare those rankings with Year 5 T1Lev ratio stress test projections. As shown in Table VI, the rank correlation coefficient of failure probability with the 2007 simulation is 0.51, and the correlation coefficient with the 2017 simulation is 0.41.

Finally, we compare T1Lev ratio rankings in Year 5 with CRE rankings—banks ranked by their proportion of CRE loans to total loans. Because the recession hit CRE loans particularly hard, we might expect a strong correlation between banks with high CRE loan concentrations and banks with poor performance in the stress tests. Again, banks are ranked from highest to lowest risk. The spearman rank correlation coefficient is 0.40 for the 2007 simulation and 0.16 for the 2017 simulation, much lower than the other correlations.

In sum, projections from our stress test model built on historical loss rates are highly correlated in sample during the 2008-2012 period with banks that failed or had early warning indicators with heightened risk in 2007. Depending on the capital threshold, stress test projections identify between

[18] The variables in the early SEER model and DFP are the log of total assets, ROA, equity to assets, other real estate owned to assets, loans 30-89 days past due to assets, loans 90 or more days past due to assets, nonaccrual loans to assets, securities to assets, and jumbo CDs to assets. Interestingly, this model performed better than a model estimated on bank failures between 2006 and 2009.

54% and 74% of the banks that failed between 2008 and 2012, and the projections are correlated with banks that had relatively low T1Lev ratios and high failure probabilities in 2007.

VI. Benefits from Loan Portfolio Reallocations

In this section, we use our stress tests to quantify the benefits to insolvency risk from loan portfolio reallocations. Exposure to commercial real estate (CRE) loans increased sharply at community banks between 1990 and 2006, which prompted supervisors to issue guidance that defined CRE concentration thresholds and encouraged banks to stress test loan portfolios (Board of Governors. 2006.) Figure VI shows that CRE lending as a percent of total loans more than doubled from 23% in 1991 to 50% in 2007. NFR and CLD loans grew the fastest, while farmland (FRM) and multifamily (MFM) remained relatively small portions of CRE loans throughout the period. Interestingly, CRE concentration has remained high since the financial crisis pinnacle in 2008. Even as late as 2017, CRE lending comprised 48.6% of total loans.

The financial crisis and subsequent recession revealed the substantial risk to community banks resulting from high CRE concentrations. Indeed, 23% of banks that exceeded both CRE thresholds established in the 2006 guidance failed during the ensuing economic downturn (Friend, Glenos and Nichols, 2013). Banks with high concentrations of CLD loans were particularly vulnerable to the economic downturn and collapse of real estate prices. Figure VII plots mean chargeoff rates by CRE loan type for community banks between 2008 and 2017. Of the four categories, CLD incurred the highest chargeoffs. The mean chargeoff rate for CLD in 2009 was 2.7%; in contrast, mean chargeoffs for NFR loans never exceeded 0.6%.

At the same time the CRE guidance was finalized, changes to the Call Report were introduced that refined CRE loan categories. Beginning in 2007 (and finalized in 2008), the Call Reports separated nonfarm nonresidential loans into owner-occupied (NFR-OWN) and "other" non-owner occupied (NFR-OTH) loans. It also split construction and land development loans into 1-4 family construction loans (CLD-RES) and "other" construction loans (CLD-OTH). Appendix C describes these changes in detail. The presumption by bankers and regulators was that owner-occupied properties would be relatively less risky because the tenants had more skin in the game. Similarly, residential construction loans were presumably less risky than other construction loans because defaults on residential construction were historically low (Federal Register, 2005). As shown in Figure VIII, between 2007 and 2017, an average 47% of NFR loans were owner occupied, and 25% of CLD loans were residential construction. Since 2012, the share of NFR loans that are owner occupied has decreased, and the share of CLD loans that are residential has increased.

We use the community bank stress test to assess risk-reduction benefits from hypothetical loan portfolio adjustments. We focus initially on portfolio adjustments between residential CLD loans (CLD-RES) and "other" CLD loans (CLD-OTH), and between owner-occupied NFR (NFR-OWN) loans and "other" NFR (NFR-OTH) loans. Figure IX plots mean chargeoff rates by these four loan categories for all U.S. community banks between 2007 and 2017. Chargeoff rates for residential construction loans were lower than other construction loans after 2009, though the chargeoff rates were similar before then. In addition, chargeoff rates for owner-occupied NFR loans were slightly lower than for other NFR loans for most the period. More importantly, defaults on NFR loans were much lower than defaults on CLD loans. These patterns suggest that community banks can achieve significant risk reduction from shifting lending from CLD to NFR rather than shifting lending within CLD and NFR.

We construct five hypothetical balance sheets for the 7,125 U.S. community banks at year-end 2007, each time restarting with the actual 2007 data so that the changes are not cumulative. We first place all CLD-OTH loans into the CLD-RES category and run the stress test. We then transfer all CLD-OTH loans into CLD-RES. We repeat the exercise for NFR loans, placing all of them in NFR-OWN and NFR-OTH, respectively. Finally, we shift all CLD loans to NFR loans by jointly transferring all CLD-RES loans into NFR-OWN, and all CLD-OTH loans into NFR-OTH. In all, we create five distinct datasets with hypothetical loan portfolios using 2007 as Y0 for the stress tests.

Stress test results in Table VII show that just one portfolio reallocation results in large differences in the number of banks that cross the capital thresholds relative to the base case. Shifting loans from CLD to NFR as shown in Row 6 of the table leads to a reduction from 762 to 262 in the number of banks with a T1Lev ratio below 2%, and a reduction from 1638 to 925 in the number of banks with a T1Lev ratio below 6%. Of course, this hypothetical loan reallocation is an extreme example where banks make no construction loans, but even modest shifts from CLD into NFR bring risk reduction benefits.

Table VII also shows that within-NFR portfolio reallocation from NFR-OTH into NFR-OWN (Row 4) results in a modest reduction in the number of banks that cross the capital thresholds, and reallocation from NFR-OWN to NFR-OTH (Row 5) results in a modest increase. This result is consistent with Figure IX that shows slightly lower chargeoff rates for NFR-OWN. For within-CLD loan reallocations, however, we observe unexpected results. The shift from CLD-OTH to CLD-RES (Row 2) increases the number of banks with less than 2% capital, perhaps because chargeoff rates for CLD-RES are slightly higher than chargeoff rates for CLD-RES in 2008 and 2009. Nevertheless, portfolio reallocations within CLD and NFR lead to modest changes in stress test outcomes relative to portfolio shifts from CLD to NFR.

VII. Conclusion

We develop an historical loss macro stress test that can be used by U.S. community banks and supervisors. Each bank undergoes a severely adverse five-year scenario where the bank experiences a chargeoff rate on a given loan type equal to the 90^{th} percentile chargeoff rate experienced by community banks in its geographical market each year between 2008 and 2012. The model naturally captures tail risk and avoids model error inherent in econometric approaches. We show that it more accurately projects patterns of actual community banks distress in the years surrounding the financial crisis and Great Recession.

More than ten years after the 2007-2009 financial crisis, community banks are well prepared to weather a similar shock. Stress tests results beginning with bank condition in 2017 show significantly lower insolvency risk than stress tests run on community banks in 2007 primarily because banks in 2017 have much lower concentrations in construction and land development loans. The Tax Cut and Jobs Act of 2017, however, offset some of these improvements for banks in 2018 because it weakens the automatic stabilizer effect from net operating losses and makes it more difficult for banks to convert accrued tax benefits to Tier 1 capital. Finally, loan portfolio diversification within each of the construction land development and nonfarm nonresidential loan categories results in little improvement in stress test outcomes because chargeoff rates within each of those categories are similar, but replacing construction land development loans with nonfarm nonresidential loans can lead to a substantial reduction in bank insolvency risk from a severely adverse shock.

References

Agarwal, S., Hauswald, R. (2010). Distance and Private Information in Lending. The Review of Financial Studies, 23(7): 2757–2788.

Barth, J., Han, S., Joo, S., Bok Lee, K., Maglic, S., Shen, X. (2018). Forecasting net charge-off rates of banks: What model works best? Quantitative Finance and Economics, 2(3), 554-589. doi:10.3934/QFE.2018.3.554

Board of Governors of the Federal Reserve System. (2006). Concentrations in Commercial Real Estate Lending, Sound Risk Management Practices, December. Joint final guidance. Docket No. OP-1248. Accessed 19 January 2020 at https://www.federalreserve.gov/newsevents/pressreleases/bcreg20061206a.htm.

Board of Governors of the Federal Reserve System. (2012). Statement to Clarify Supervisory Expectations for Stress Testing by Community Banks, May. Accessed 19 January 2000 at https://www.federalreserve.gov/newsevents/pressreleases/bcreg20120514b.htm.

Board of Governors of the Federal Reserve System. (2017). Dodd-Frank Act Stress Test 2019: Supervisory Stress Test Methodology. Accessed 19 January 2020 at https://www.federalreserve.gov/publications/2017-june-dodd-frank-act-stress-test-preface.htm.

Board of Governors of the Federal Reserve System. (2019). Dodd-Frank Act Stress Test 2019: Supervisory Stress Test Methodology. Accessed 19 January 2020 at https://www.federalreserve.gov/publications/files/2019-march-supervisory-stress-test-methodology.pdf

Cole, R., Cornyn, B., Gunther, J. (1995). FIMS: A New Monitoring System for Banking Institutions. Federal Reserve Bulletin 81, 1-15.

Cole, R., White. L. (2017). When time is not on our side: The costs of regulatory forbearance in the closure of insolvent banks. Journal of Banking and Finance 80: 235–249.

Covas, F., Rump, B., & Zakrajšek, E. (2014). Stress-testing US bank holding companies: A dynamic panel quantile regression approach. International Journal of Forecasting 30:691-713.

Covas, F. (2017). The Capital Allocation Inherent in the Federal Reserve's Capital Stress Test. The Clearing House. Accessed 10 December 2018 at https://www.theclearinghouse.org/-/media/tch/documents/tch%20weekly/2017/20170130_tch_research_note_implicit_risk_weights_in_ccar-final.pdf.

DeYoung, R., & Fairchild, J. (2018) Stress testing community banks. Working paper. Accessed 19 January 2020 at https://business.ku.edu/sites/business.ku.edu/files/images/general/Research/stress_v8.3_class_fed.pdf

Federal Deposit Insurance Corporate. (2018) FDIC Small Business lending Survey. Accessed 12 July 2019 at https://www.fdic.gov/bank/historical/sbls.

Federal Financial Institutions Examination Council. (2006). Revisions to the Reports of Condition and Income (Call Report). Financial Institutions Letter FIL-7-2006. January. Accessed 19 January 2000 at https://www.ffiec.gov/whatsnew_200601.htm.

Federal Register. (2005). Proposed Agency Information Collection Activities; Comment Request, 70 (162) 49363-49372.

Federal Register (2019). Amendments to Policy Statement on the Scenario Design Framework for Stress Testing. 28 February: 84(40) 6654.

Friend, K. & Glenos, H. & Nichols J. B. An analysis of the impact of the commercial real estate concentration guidance. Office of the Comptroller of the Currency. Accessed 30 March 2015 at http://www.federalreserve.gov/bankinforeg/cre-20130403a.pdf.

Garcia, G. (2010) Failing prompt corrective action, Journal of Banking Regulation 11:171-190.

Government Accountability Office (2011) Bank Regulation: Modified Prompt Corrective Action Framework Would Improve Effectiveness. GAO-11-612. Accessed 19 January 2020 at https://www.gao.gov/products/GAO-11-612.

Grover, S., McCracken, M. W. (2014). Factor-based prediction of industry-wide bank stress. Federal Reserve Bank of St. Louis Review, 96(2), 173.

Guerrieri, L., & Welch, M. (2012). Can macro variables used in stress testing forecast the performance of banks? Washington, DC: Div. of Research & Statistics and Monetary Affairs, Federal Reserve Board. Accessed 19 January 2020 at https://www.federalreserve.gov/pubs/feds/2012/201249/201249pap.pdf.

Hall, J., Kern, D., King, T., Lee, K. Yeager, T. (2011). A Value-at-Risk Approach to Commercial Real Estate Portfolio Stress Testing at U.S. Community Banks, Journal of Risk Management in Financial Institutions 5.1 (2011): 60-75.

Hirtle, B., Kovner, A., Vickery, J., & Bhanot, M. (2016). Assessing financial stability: The capital and loss assessment under stress scenarios (CLASS) model. Journal of Banking and Finance, 69, S35-55.

Jacobs, M. (2016). Stress testing and a comparison of alternative methodologies for scenario generation. Journal of Applied Finance and Banking, 6(6), 1-7.

Kapinos, P., & Mitnik, O. A. (2016). A top-down approach to stress-testing banks. Journal of Financial Services Research, 49(2), 229-264.

Miller, S., Olson, E., Yeager, T. (2015) The relative contributions of equity and subordinated debt signals as predictors of bank distress during the financial crisis. Journal of Financial Stability: 16: 118-137.

Petersen, M., & Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. Journal of Finance, 57(6), 2533-2570.

Schmieder, C., Puhr, C., Hasan, M. (2014) Next-Generation Applied Solvency Stress Testing in A Guide to IMF Stress Testing: Methods and Models / editor, Ong, L. L., International Monetary Fund, Washington D.C.: 59-69. Accessed 19 January 2020 at https://www.elibrary.imf.org/fileasset/misc/toolkit/pdf/chap5.pdf?redirect=true.

Wagner, A. F., R. J. Zeckhauser, A. Ziegler. (2018). Company stock price reactions to the 2016 election shock: Trump, taxes, and trade. Journal of Financial Economics, 130(2): 428-451.

Table I: Mean 90th percentile chargeoff rates across U.S. community bank markets

Table displays mean 90th percentile net chargeoff rates across U.S. community bank markets for the years 2008-2012. Panel A lists values for rural banks (headquartered in counties not in MSAs), and Panel B lists values for urban banks (headquartered in MSAs). Call Reports specify six types of commercial real estate (*CRE*) loans. Multifamily (*MFM*) loans are secured by properties with five or more units. Nonfarm nonresidential (*NFR*) loans are secured by business real estate and they are divided into other (*NLR-OTH*) and owner-occupied (*NFR-OWN*). Farmland (*FRM*) loans are secured by farm real estate. Construction and land development (*CLD*) loans are divided into residential (*CLD-RES*) and other (*CLD-OTH*). Call Reports specify five other loan types. Consumer (*CN*) loans are to individuals for items such as automobiles and credit cards loans. Mortgage (*MTG*) loans are residential real estate loans secured by property with less than five units. Commercial and Industrial (*C&I*) loans are business loans not secured by real estate. Agricultural (*AG*) loans are loans for agricultural production not secured by farmland. We define *Other* loans as all other loan types.

Panel A. Rural banks

	Loan type	2008	2009	2010	2011	2012
CRE	MFM	0.00%	0.00%	0.22%	0.00%	0.00%
	NFR-OTH	0.45%	1.35%	1.81%	1.83%	1.18%
	NFR-OWN	0.32%	1.11%	1.25%	1.21%	0.94%
	FRM	0.00%	0.00%	0.02%	0.03%	0.00%
	CLD-OTH	1.70%	6.90%	8.33%	6.92%	3.95%
	CLD-RES	0.99%	3.48%	3.19%	1.31%	0.15%
	Consumer	1.80%	2.17%	1.87%	1.52%	1.39%
	Mortgages	0.49%	0.89%	1.13%	1.07%	0.83%
	C&I	2.57%	4.26%	3.58%	3.02%	2.04%
	Agriculture	0.13%	0.38%	0.25%	0.10%	0.02%
	Other	3.43%	3.65%	2.79%	1.99%	2.17%

Panel B. Urban banks

	Loan type	2008	2009	2010	2011	2012
CRE	MFM	0.00%	1.19%	2.36%	1.78%	0.78%
	NFR-OTH	0.44%	1.99%	2.55%	2.42%	1.67%
	NFR-OWN	0.36%	1.19%	1.73%	1.68%	1.29%
	FRM	0.00%	0.00%	0.09%	0.03%	0.01%
	CLD-OTH	3.88%	10.73%	12.02%	9.46%	5.49%
	CLD-RES	4.85%	11.15%	10.02%	6.73%	2.03%
	Consumer	2.86%	3.81%	3.74%	2.81%	2.15%
	Mortgages	1.05%	2.65%	2.64%	2.32%	1.68%
	C&I	2.71%	5.85%	5.49%	3.83%	2.85%
	Agriculture	0.07%	0.18%	0.20%	0.03%	0.00%
	Other	3.43%	7.66%	5.81%	3.14%	2.37%

Table II: Stress test results for U.S. community banks, 2017-2022

Summary statistics of stress-test results for the sample of 4,846 community banks based on their financial conditions at year-end 2017 (Y0). *T1Lev*<2%, *T1Lev* <6%, and *ROA*<0% are the number of banks each year with projected Tier 1 leverage ratios and return on assets, respectively, below the threshold.

	Year 0	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**
T1Lev Ratio	2017	2018	2019	2020	2021	2022
Mean	11.46%	11.42%	11.02%	10.61%	10.30%	10.17%
Median	10.42%	10.45%	10.16%	9.94%	9.78%	9.74%
Min	2.66%	-64.92%	-141.77%	-218.17%	-294.09%	-369.51%
Max	98.75%	100.50%	102.88%	105.18%	110.65%	116.97%
StdDev	5.85%	5.93%	6.39%	7.08%	7.87%	8.71%
T1Lev< 2%	0	7	21	49	97	153
T1Lev<6%	31	44	104	295	474	563
Chargeoffs to Loans	2017	2018	2019	2020	2021	2022
Mean	0.18%	1.03%	2.23%	2.27%	1.88%	1.31%
Median	0.05%	0.85%	1.63%	1.71%	1.43%	1.04%
Min	-1.75%	0.05%	0.04%	0.09%	0.04%	0.02%
Max	38.80%	6.19%	15.10%	16.06%	16.56%	14.78%
StdDev	0.82%	0.72%	1.71%	1.76%	1.43%	1.01%
ROA	2017	2018	2019	2020	2021	2022
Mean	0.97%	0.49%	-0.08%	-0.09%	0.06%	0.30%
Median	0.97%	0.50%	0.10%	0.06%	0.18%	0.36%
Min	-174.42%	-130.32%	-130.32%	-130.32%	-130.32%	-130.32%
Max	45.28%	32.56%	32.56%	32.56%	32.56%	32.56%
StdDev	2.88%	2.18%	2.26%	2.26%	2.21%	2.17%
ROA<0%	261	730	2132	2246	1908	1246

Table III: Comparison of 2017 community bank stress test with DFAST

Table compares cumulative nine-quarter loan loss rates from the 2017 community bank stress test results for banks headquartered in urban markets with the 2017 Dodd Frank Stress Test (DFAST) adverse scenario results from Table VII of the report. We compute the community bank loss rates using the same procedure as DFAST where loss rates are accumulated over a nine-quarter horizon and divided by the average loan balances over the same period. Consumer loans under DFAST include only the "other consumer" category. n/a signifies that the loss rate is not available.

	Cumulative Nine-Quarter Loss Rates	
Loan type	Urban Community Banks	DFAST
CRE	4.6%	7.0%
Consumer	7.6%	5.9%
Mortgages	4.4%	3.4%
C&I	9.9%	6.4%
Agriculture	0.3%	n/a
Other	12.5%	n/a
Portfolio Loss Rate	5.0%	5.8%

Table IV: Stress test results for U.S. community banks, 2007-2012

Summary statistics of stress-test results for the sample of 7,125 community banks based on their financial conditions at year-end 2007 (Y0). *T1Lev*<2%, *T1Lev* <6%, and *ROA*<0% are the number of banks each year with projected Tier 1 leverage ratios and return on assets, respectively, below the threshold.

	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
T1Lev Ratio	2007	2008	2009	2010	2011	2012
Mean	12.66%	11.60%	10.86%	10.11%	9.57%	9.35%
Median	9.67%	9.67%	9.28%	8.97%	8.83%	8.82%
Min	2.23%	-4.14%	-13.02%	-24.85%	-38.79%	-68.83%
Max	1811.14%	112.35%	134.95%	156.82%	178.00%	198.50%
StdDev	25.03%	7.81%	7.86%	8.21%	8.62%	8.95%
T1Lev< 2%	0	50	124	424	648	762
T1Lev<6%	44	146	674	1197	1536	1638
Chargeoffs to Loans	2007	2008	2009	2010	2011	2012
Mean	0.24%	1.47%	3.30%	3.29%	2.63%	1.69%
Median	0.09%	1.05%	2.09%	2.18%	1.79%	1.26%
Min	-20.41%	0.00%	0.00%	0.10%	0.08%	0.04%
Max	13.58%	41.33%	35.26%	93.03%	31.95%	17.56%
StdDev	0.66%	1.33%	3.07%	3.17%	2.31%	1.38%
ROA	2007	2008	2009	2010	2011	2012
Mean	1.08%	0.42%	-0.46%	-0.45%	-0.16%	0.25%
Median	1.13%	0.51%	0.00%	-0.03%	0.11%	0.36%
Min	-36.30%	-23.95%	-24.48%	-48.47%	-24.35%	-24.14%
Max	60.52%	62.83%	62.83%	62.83%	62.83%	62.83%
StdDev	1.89%	1.74%	2.21%	2.23%	1.95%	1.73%
ROA<0%	652	1705	3589	3672	3276	2333

Table V: Comparison of hypothetical stress test outcomes

Table displays the percent of community banks with projected Tier 1 Leverage (T1Lev) and equity-to-asset ratios below the respective threshold. Indented rows beneath the baseline in each panel are stress test results relative to the baseline. Stress tests "under TCJA" include effects from the Tax Cuts and Jobs Act of 2017. *N* is the number of banks in the simulation.

Row	Stress Test Description	T1Lev <2%	T1Lev <6%	Equity < 2%	Equity < 6%	N
1	Baseline 2007 stress test	10.7%	23.0%	9.5%	20.6%	7125
2	with banks that did not exist in 2017	19.5%	35.4%	16.8%	31.3%	2509
3	with banks that also existed in 2017	4.9%	14.7%	4.4%	13.4%	4616
4	Baseline 2017 stress test	3.2%	11.6%	2.7%	11.0%	4846
5	with banks that also existed in 2007	2.5%	10.4%	2.3%	9.9%	4616
6	with 2007 loan shares	4.5%	14.3%	4.0%	13.7%	4616
7	under TCJA	6.8%	16.7%	4.5%	14.0%	4846
8	Baseline 2018 stress test under TCJA	5.9%	14.7%	3.6%	11.9%	4598

Table VI: Spearman rank correlations of early warning signals and projected capital

Table presents Spearman rank correlations of the actual ranks of each of three early warning signals in Y0 with the projected stress-test Tier 1 Leverage ratio in Y5. Each early warning signal is ranked from highest risk to lowest risk. The correlations in the first column are from 2007 stress tests where Y0 is 2007 and Y5 is 2012, and those in the second column are from 2017 stress tests where Y0 is 2017 and Y5 is 2022. *T1Lev Ratio* is the Tier Leverage ratio, *Failure probability* is a logit model that estimates the likelihood of bank failure in the following two years, and *CRE to assets* is the ratio of commercial real estate loans to assets.

Actual rank of early warning signal in Y0	Correlation with projected T1Lev Ratio rank in Y5	
	2007 Stress Test	2017 Stress Test
T1Lev Ratio	0.64	0.75
Failure probability	0.51	0.41
CRE to assets	0.40	0.16

Table VII: Stress test outcomes from hypothetical loan portfolio shifts

Table displays the number of community banks in 2007 with projected stress-test T1Lev ratios below the respective 2% and 6% thresholds. Row 1 results are the baseline 2007 stress tests. The remaining rows reflect hypothetical stress tests that shift loans from the first loan type into the second loan type. For example, Row 2 reports simulation results after shifting all *CLD-OTH* loans into *CLD-RES* loans. *CLD-OTH* and *CLD-RES* are, respectively, other and residential construction and land development loans. *NFR-OTH* and *NFR-RES* are, respectively, other and residential nonfarm nonresidential commercial real estate loans.

Portfolio Shifts	T1Lev <2%	T1Lev <6%
1. Actual 2007 Loan Portfolio	762	1638
2. CLD-OTH → CLD-RES	800	1570
3. CLD-RES → CLD-OTH	754	1672
4. NFR-OTH → NFR- OWN	717	1577
5. NFR-OWN → NFR-OTH	810	1697
6. CLD-RES → NFR- OWN *and* CLD-OTH → NFR-OTH	262	925

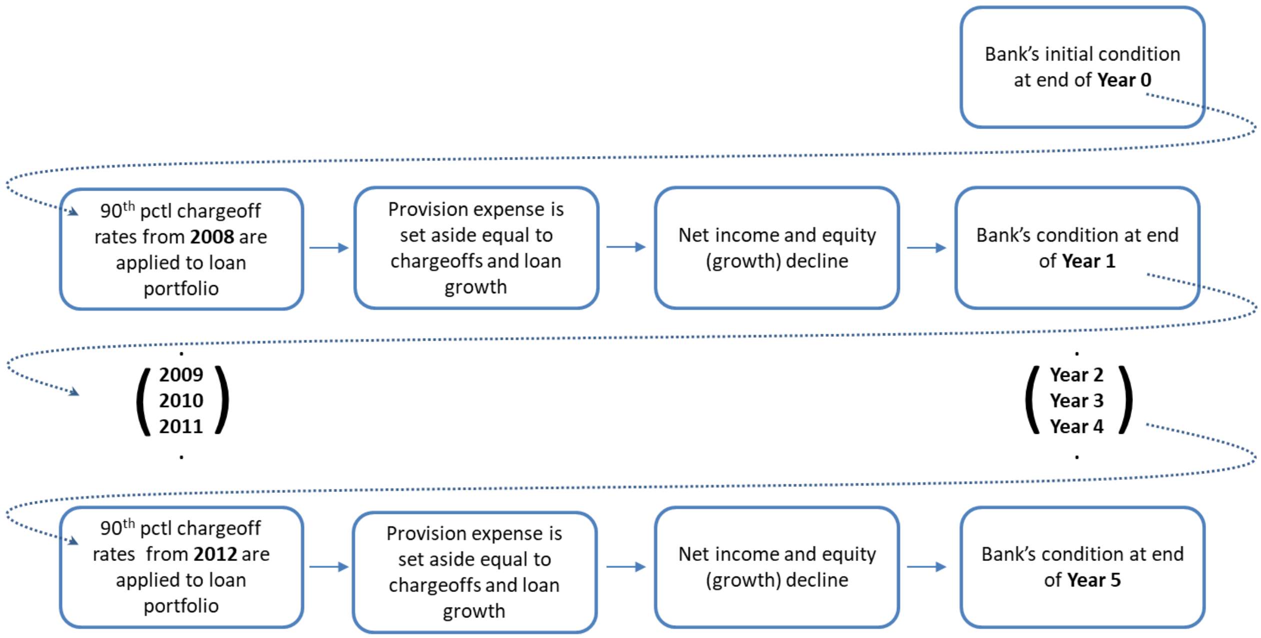


Figure I: Five-Year Simulation Flow Chart
Flow chart summarizes the logic of the five-year community bank stress test. From the initial condition at Y0, the bank incurs in Y1 the 90th percentile loan chargeoff rates of its market from 2008 and sets aside provision expense equal to the net chargeoffs plus realized loan growth targeted at 3%. After-tax net income (35% tax rate) not paid as dividends (Min($0, Payout Ratio$_{2007}$)) add to or subtract from the bank's retained earnings and capital. The bank then incurs the 90th percentile chargeoff rates from 2009 in Y2, and the pattern repeats through Y5.

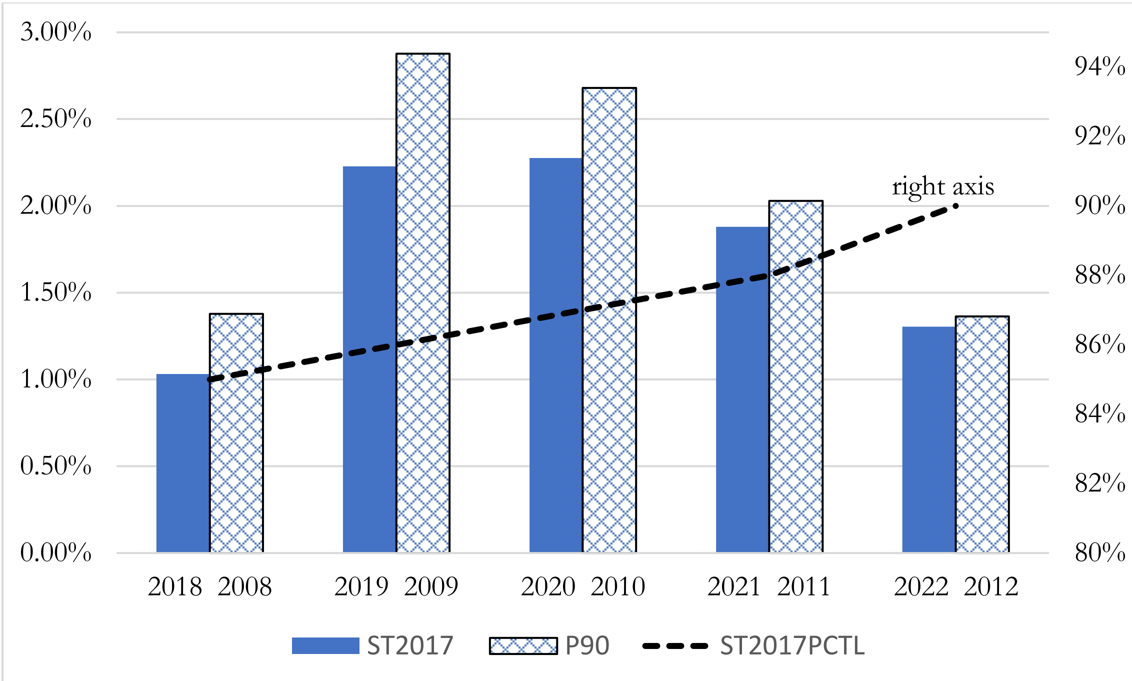


Figure II: Actual and Simulated Portfolio Chargeoff Rates
Figure compares means of community bank projected net chargeoff rates with actual net chargeoff rates. *ST2017* plots the mean projected net chargeoff rate for each year 2018-2022 based on banks' initial conditions at year-end 2017. *P90* plots the mean actual 90th percentile net chargeoff rate across all markets for each year 2008-2012. *ST2017PCTL* (right axis) plots the percentile of the projected net chargeoff for the years 2018-2022 relative to actual mean 90th percentile net chargeoff rates from the years 2008-2012.

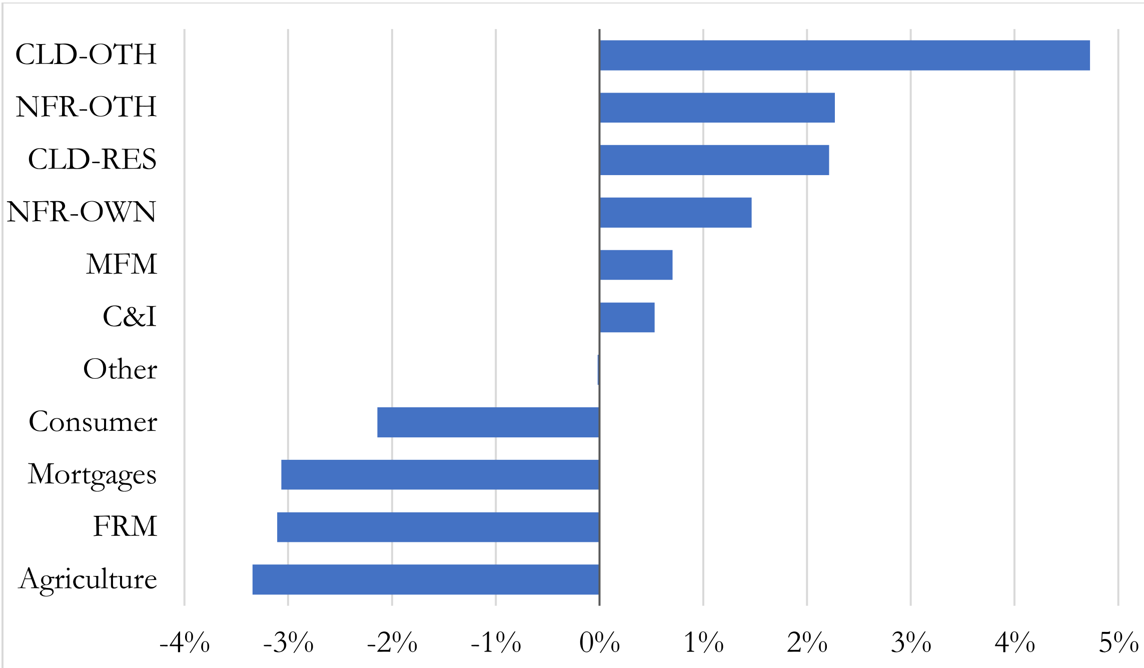
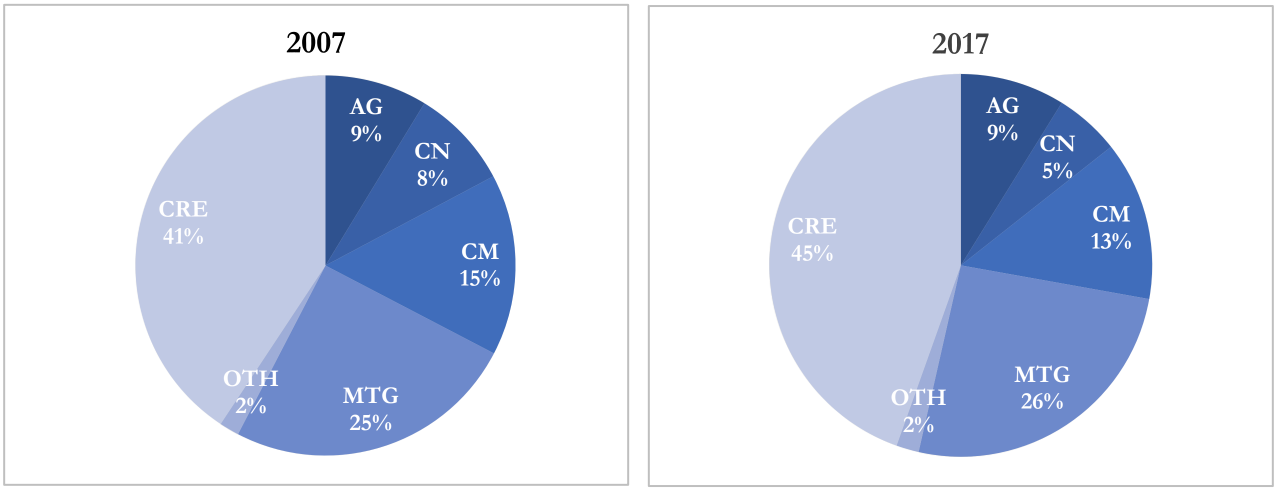
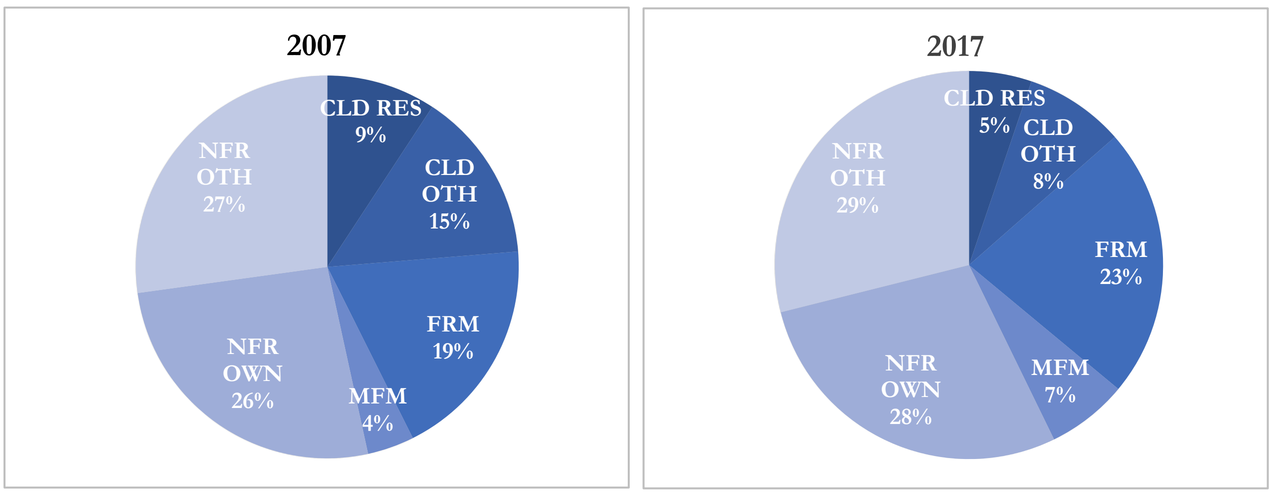


Figure III: Differences in mean loan shares in 2007 by bank existence in 2017
Figure plots the differences in mean loan shares (loan category amount scaled by total loans) in 2007 by bank status in 2017. Differences are computed as mean loan share of banks that did not exist in 2017 less banks that did exist in 2017. See Table I for loan category definitions.

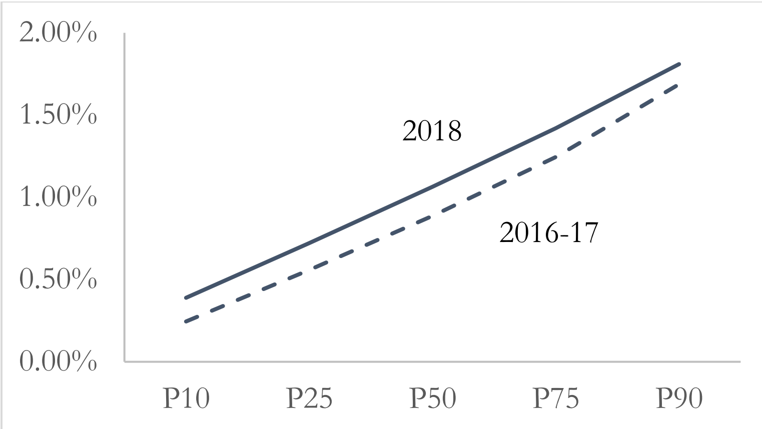


Panel A. Major Loan Categories

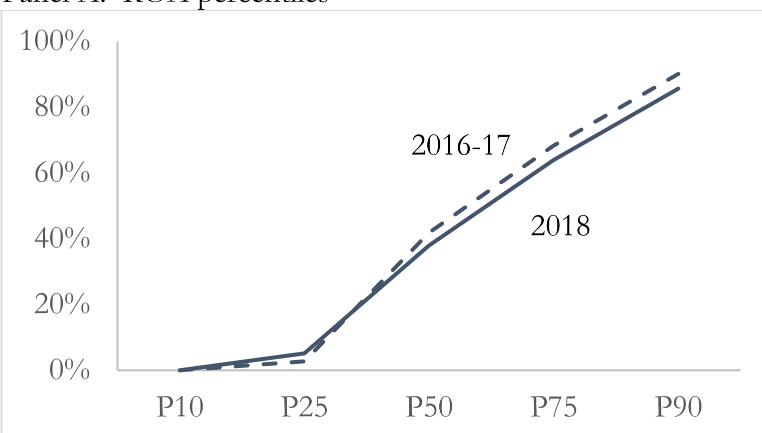


Panel B. Commercial Real Estate Categories

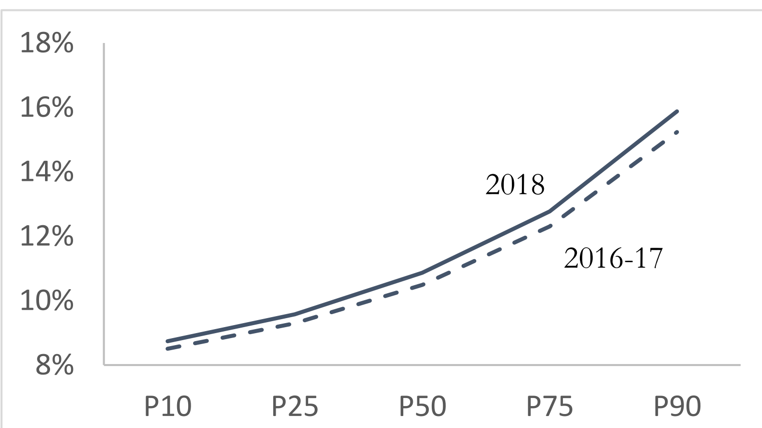
Figure IV: Mean Community Bank Loan Portfolio, 2007 and 2017
Figures in Panel A plot mean loan share by major loan category for all community banks that existed in both 2007 and 2017. *CRE* is commercial real estate loans, *AG* is agricultural loans, *CN* is consumer loans, *CM* is commercial and industrial loans, *MTG* is residential mortgage loans, and *OTH* is all other loans. The left-hand chart plots loan shares at year-end 2007, and the right-hand chart, at year-end 2017. Figures in Panel B plot for all community banks that existed in both 2007 and 2017 the loan share by CRE category. *CLD-RES* and *CLD-OTH* are, respectively, residential and other construction and land development loans, *FRM* is loans secured by farmland, *MFM* is multifamily loans, and *NFR-OWN* and *NFR-OTH* are, respectively, owner-occupied and other nonfarm nonresidential loans. The left-hand chart plots CRE loan shares at year-end 2007, and the right-hand chart, at year-end 2017.



Panel A. ROA percentiles



Panel B. Dividend payout percentiles



Panel C. T1Lev ratio percentiles

Figure V: Bank earnings, dividend payouts, and capital ratios

Figure plots the 10th through 90th percentiles for select community bank ratios. Ratios for 2016-17 are the average of the two years. Panel A plots return on assets (ROA); Panel B, the dividend payout ratio; and Panel C, the Tier 1 leverage ratio.

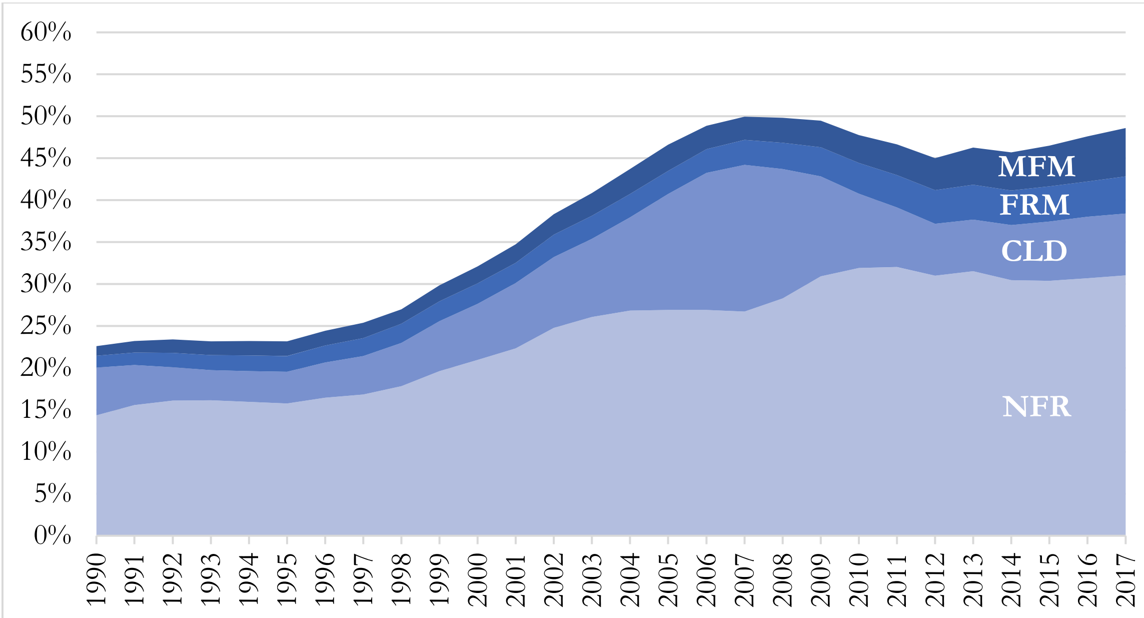


Figure VI: CRE Loan Concentration at Community Banks 1990-2017
Figure plots for all community banks between 1990 and 2017 asset-weighted loan shares of the four primary commercial real estate loan categories: nonfarm nonresidential (*NFR*), construction and land development (*CLD*), farmland (*FRM*), and multifamily (*MFM*).

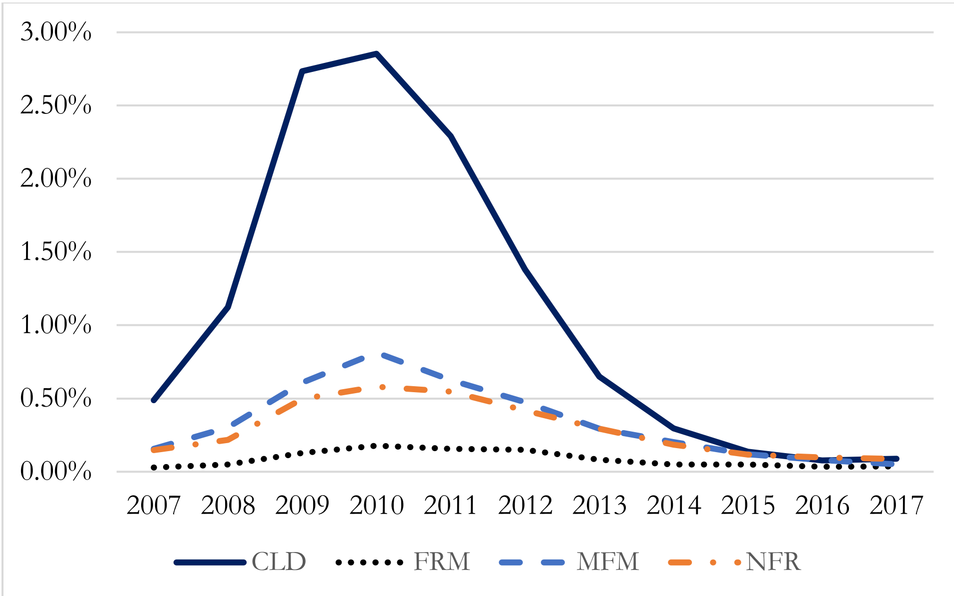


Figure VII: Mean Chargeoff Rates by CRE Loan Type
Figure plots community bank mean net chargeoff rates between 2007 and 2017 for construction and land development (*CLD*), farmland (*FRM*), multifamily (*MFM*), and nonfarm nonresidential (*NFR*).

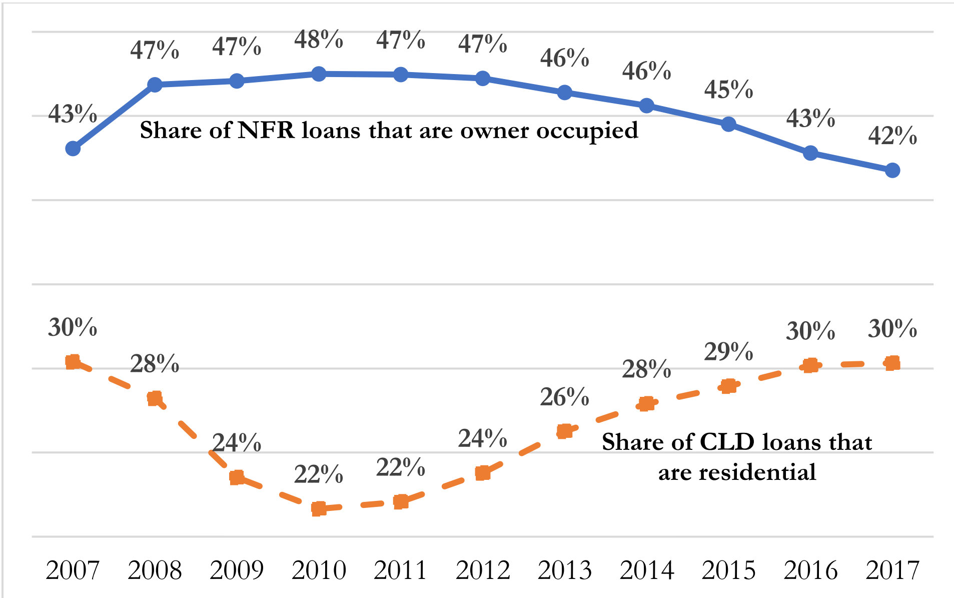


Figure VIII: Relative Shares of CLD and NFR Loans
Figure plots for all community banks from 2007-2017 the share of total nonfarm nonresidential loans that are owner occupied (NFR-OWN), and the share of total construction and land development loans that are residential (CLD-RES).

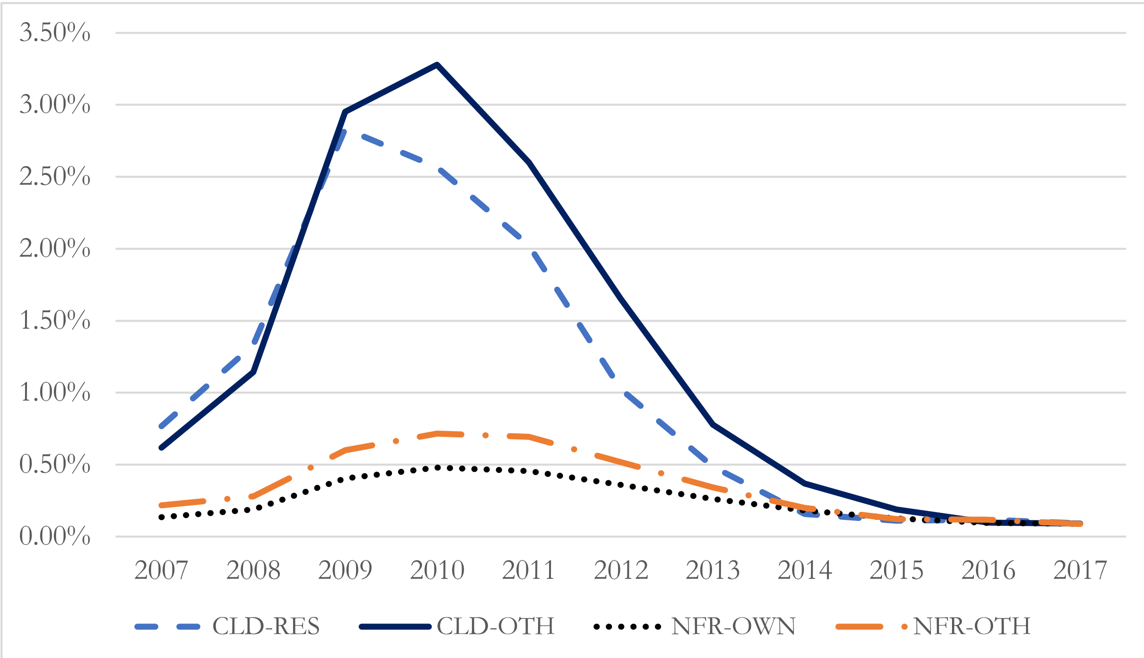


Figure IX: Mean Chargeoff Rates for CLD and NFR Subcategories
Figure plots community bank mean net chargeoff rates between 2007 and 2017 for residential (CLD-RES) and 'other' (CLD-OTH) construction and land development loans. It also plots net chargeoff rates for owner-occupied (NFR-OWN) and 'other' (NFR-OTH) nonfarm nonresidential loans.

Appendix Table AI: Inputs for Community Bank Stress Test

Microsoft Excel worksheet with the required Call Report inputs to run the stress test.

Call Report Date: Year 0		Metropolitan Statistical Area	
Year:	2017	myMSA	
Quarter:	4		
Cert	COHORT	PCTL Loss Rate	Asset Growth Rate
1	MSA	90%	3.0%

Enter dollar amounts as year-to-date

Commercial Real Estate	Loan Amount ($000s)	Annual Interest Rate (%)	YTD Net Losses ($000s)
MULTIFAM	57,296	3.50%	69
NFR-Other	143,979	3.50%	755
NFR-Owner Occupied	178,590	3.50%	154
FARM	0	3.50%	0
CLD-Other	36,299	3.50%	113
CLD-Residential	7,875	3.50%	33

Other Loans & Securities	Asset Amount ($000s)	Annual Interest Rate (%)	YTD Net Losses ($000s)
Mortgages	198,888	6.33%	1,639
Consumer	233,220	2.74%	437
Commercial & Industrial	70,071	2.99%	2
Agricultural	0	2.99%	0
Other Loans	7,453	6.33%	0
Securities	217,572	2.44%	
Federal Funds Sold	0	0.03%	
Interest Bearing Balances	6,215	0.02%	

Other Items	$000s
Interest expense	4,789
Noninterest expense	35,090
Noninterest income	12,999
Provision expense	3,113
Securities & Extra. gains	-5
Taxes	3,379
Dividend Payout	5,300
Loan Loss Reserves (ALLL)	17,416
Average assets	1,233,734
Non-earning assets	95,637
Total Liabilities	1,106,691
Trading Assets	0
Tier 1 capital	107,498
Total assets for leverage ratio	1,210,128

Appendix Table AII. Stress Test Results for a Sample Community Bank

Microsoft Excel worksheet with the stress test output for a sample community bank.

Balance Sheet ($000s)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
Interest Bearing Balances	6,215	6,401	6,593	6,791	6,995	7,205
Federal Funds Sold	0	0	0	0	0	0
Securities	217,572	224,099	230,822	237,747	244,879	252,226
Total Loans	933,671	950,895	960,139	968,103	978,733	992,484
LLR	17,416	17,737	17,910	18,058	18,257	18,513
Net Loans	916,255	933,157	942,229	950,045	960,477	973,971
Trading Assets	0	0	0	0	0	0
Total Earning Assets	1,140,042	1,163,658	1,179,645	1,194,583	1,212,351	1,233,402
Non-Earning Assets	95,637	97,618	98,959	100,212	101,703	103,469
Total Assets	1,235,679	1,261,276	1,278,604	1,294,795	1,314,054	1,336,870
Liabilities	1,106,691	1,131,329	1,151,290	1,170,984	1,192,064	1,214,828
Equity	128,988	129,947	127,314	123,812	121,990	122,042

Net Chargeoffs (annualized in $000s)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
Net chargeoffs	3,202	10,787	19,282	20,840	18,413	15,611

Income Statement (annualized in $000s)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
Interest income	41,718	42,591	43,218	43,766	44,394	45,167
Interest expense	4,789	4,888	4,955	5,018	5,093	5,181
Net Interest Income	36,929	37,703	38,263	38,748	39,301	39,986
Noninterest expense	35,090	35,817	36,309	36,769	37,316	37,964
Noninterest income	12,999	13,268	13,451	13,621	13,823	14,064
Provision	3,113	11,108	19,455	20,989	18,611	15,868
Securities & Extraordinary gains	-5	0	0	0	0	0
Operating income	11,720	4,046	-4,050	-5,389	-2,802	218
Taxes	3,379	1,416	-1,418	-1,886	-981	76
Net income	8,341	2,630	-2,633	-3,503	-1,821	142
Dividend Payout	5,300	1,671	0	0	0	90
Retained Earnings	3,041	959	-2,633	-3,503	-1,821	52

Annualized Net Chargeoffs to Loans (%)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
Commercial Real Estate	0.27%	0.73%	1.54%	1.79%	2.16%	1.16%
Multifamily	0.12%	0.77%	0.59%	0.84%	3.30%	0.68%
NFR-Other	0.52%	0.25%	2.66%	1.85%	1.56%	1.22%
NFR-Owner Occupied	0.09%	1.25%	0.99%	1.26%	0.97%	0.91%
Farm	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%
CLD-Other	0.31%	0.00%	1.25%	5.70%	8.00%	3.14%
CLD-Residential	0.42%	0.34%	0.60%	0.00%	3.17%	0.32%
Residential Mortgages	0.82%	0.63%	0.78%	1.39%	1.26%	0.98%
Consumer	0.19%	1.95%	3.38%	3.00%	1.86%	2.64%
Commercial & Industrial	0.00%	1.95%	3.44%	3.52%	1.86%	1.99%
Agriculture	0.00%	0.66%	0.36%	0.09%	0.02%	0.05%
Other Loans	0.00%	3.49%	1.88%	1.24%	1.66%	3.32%
Net chargeoffs to total loans	0.34%	1.13%	2.01%	2.15%	1.88%	1.57%

Table AII. (Cont.)

Profitability and Capital (%)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
ROA (annualized)	0.68%	0.21%	-0.21%	-0.27%	-0.14%	0.01%
ROE (annualized)	6.47%	2.02%	-2.07%	-2.83%	-1.49%	0.12%
Equity to assets	10.44%	10.30%	9.96%	9.56%	9.28%	9.13%
T1Lev ratio	8.88%	8.78%	8.45%	8.06%	7.80%	7.67%

Loans by Category ($000s)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
Commercial Real Estate	424,039	433,608	439,838	445,046	448,687	456,859
Multifamily	57,296	58,562	59,966	61,246	60,999	62,403
NFR-Other	143,979	147,930	148,321	149,938	152,022	154,677
NFR-OwnerOccupied	178,590	181,645	185,246	188,402	192,164	196,131
Farm	0	0	0	0	0	0
CLD-Other	36,299	37,388	38,029	36,937	35,001	34,920
CLD-Residential	7,875	8,083	8,276	8,524	8,501	8,729
Residential Mortgages	198,888	203,570	208,033	211,306	214,905	219,190
Consumer	233,220	235,544	234,402	234,196	236,729	237,381
Commercial & Industrial	70,071	70,764	70,378	69,937	70,697	71,371
Agriculture	0	0	0	0	0	0
Other Loans	7,453	7,409	7,488	7,617	7,716	7,683
Loan growth		1.84%	0.97%	0.83%	1.10%	1.40%

Income Statement (YTD in $000s)	*Y0*	*Y1*	*Y2*	*Y3*	*Y4*	*Y5*
Interest income	41,718	42,591	43,218	43,766	44,394	45,167
Interest expense	4,789	4,888	4,955	5,018	5,093	5,181
Net Interest Income	36,929	37,703	38,263	38,748	39,301	39,986
Noninterest expense	35,090	35,817	36,309	36,769	37,316	37,964
Noninterest income	12,999	13,268	13,451	13,621	13,823	14,064
Provision	3,113	11,108	19,455	20,989	18,611	15,868
Securities & Extraordinary gains	-5	0	0	0	0	0
Operating income	11,720	4,046	-4,050	-5,389	-2,802	218
Taxes	3,379	1,416	-1,418	-1,886	-981	76
Net income	8,341	2,630	-2,633	-3,503	-1,821	142
Dividend Payout	5,300	1,671	0	0	0	90
Retained Earnings	3,041	959	-2,633	-3,503	-1,821	52

Appendix Table AIII. Call Report Changes

The FFIEC issued FIL-7-2006 "Revisions to the Reports of Condition and Income (Call Report)" on January 27, 2006. The revisions specify that "beginning March 31, 2007, banks with $300 million or more in assets and certain banks with less than $300 million in assets will report two-way breakdowns of their real estate construction loans and their nonfarm nonresidential real estate loans in a number of Call Report schedules. All other banks with less than $300 million in assets will begin to provide these loan breakdowns as of March 31, 2008." [p. 2]

Construction and Land Development (CLD) loans were split into 1-4 family residential construction loans and *other* CLD loans. 1-4 family residential construction loans are "for the purpose of constructing 1-4 family residential properties, which will secure the loan." [p. 10]

Loans previously classified as secured by nonfarm nonresidential properties were split into loans secured by owner-occupied nonfarm nonresidential properties and loans secured by *other* nonfarm nonresidential properties. Loans secured by other nonfarm nonresidential properties are those "where the primary or a significant source of repayment is derived from rental income associated with the property (i.e., loans for which 50 percent or more of the source of repayment comes from third party, nonaffiliated, rental income) or the proceeds of the sale, refinancing, or permanent financing of the property. Thus, the primary or a significant source of repayment for 'Loans secured by owner-occupied nonfarm nonresidential properties' is the cash flow from the ongoing operations and activities conducted by the party, or an affiliate of the party, who owns the property, rather than from third party, nonaffiliated, rental income or the proceeds of the sale, refinancing, or permanent financing of the property." [p. 11]