The Impact of Prices on Analyst Cash Flow Expectations

Aditya Chaudhry*

January 12, 2024

Abstract

Analyst cash flow expectations deviate from rational expectations and correlate with stock prices. Does this correlation arise solely because these biased expectations are shared by investors who impact prices? Or do prices impact analyst expectations? I provide evidence of the latter mechanism. Using exogenous price variation from index reconstitutions and mutual fund flow-induced trading, I find analysts raise their cash flow expectations in response to price increases unrelated to fundamentals. This impact of prices on analyst cash flow expectations explains approximately half of the covariance between these objects. These results have important implications for models used to interpret analyst expectations.

*Ohio State University, Fisher College of Business. (chaudhry.127@osu.edu). This paper was previously part of my paper "Do Subjective Growth Expectations Matter for Asset Prices?" (posted on SSRN on September 3, 2022 with first draft in June 2021). I thank Niels Gormsen, Lars Hansen, Ralph Koijen, and Stefan Nagel for their invaluable advice and generous help. I think Andrea Ricciardi for excellent research assistance. I thank Anna Pavlova and Taisiya Sikorskaya for providing access to data. For comments, I thank Alexandru Barbu, Francesca Bastianello, Ricardo De La O, Xavier Gabaix, Andrei Goncalves, Valentin Haddad, Sebastian Hillenbrand, Zhiguo He, John Heaton, Lubos Pastor, Yueran Ma, Aaron Pancost, Anna Pavlova, Taisiya Sikorskaya, Alp Simsek, Pietro Veronesi, Eric Zwick, and participants of the Chicago Booth Asset Pricing Working Group, Chicago Economic Dynamics Working Group, Chicago Booth Finance Brown Bag, Inter-Finance PhD Seminar, Chicago Joint Program Conference, SFS Cavalcade North America 2022, Transatlantic Doctoral Conference 2022, 14th Annual SoFiE Post-Conference, the Machine Learning in Economics Summer Institute 2022, and the UW-Madison Junior Finance Conference. All errors are my own. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1746045. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

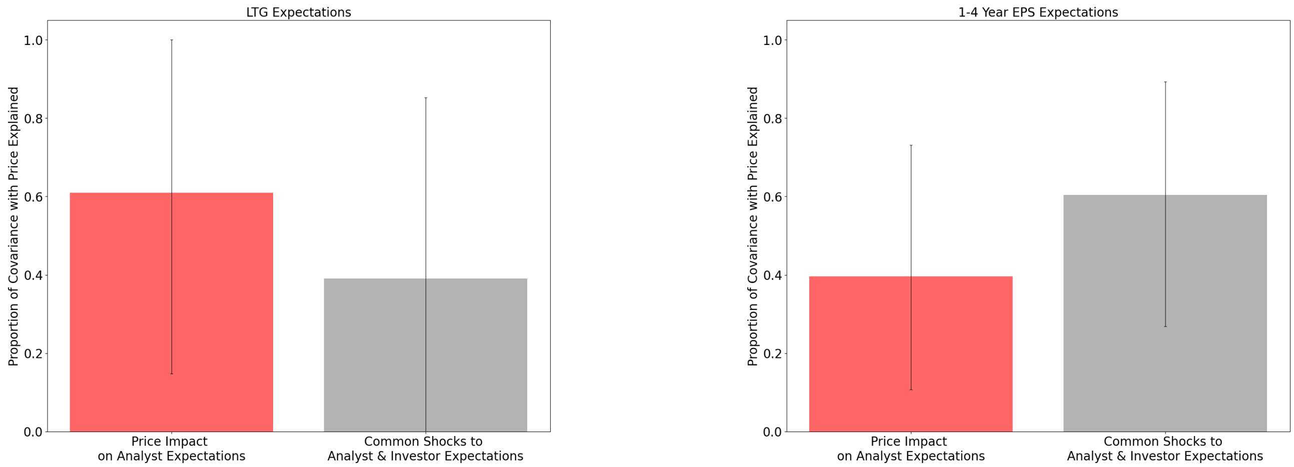
A large literature examines the impact of subjective beliefs about fundamentals — beliefs about future cash flows that may not be consistent with rational expectations — on asset prices. Due to the lack of data on investor beliefs, this literature often uses equity research analyst cash flow expectations as a proxy. Analyst cash-flow expectations correlate strongly with prices, match the magnitude of price variation, have predictable forecast errors, and negatively predict future returns. Previous work interprets these stylized facts with models that assume biased analyst cash flow expectations are shared by investors and so distort prices, but do not depend on prices (La Porta (1996); De Bondt and Thaler (1990); Bordalo et al. (2019, 2022); De La O and Myers (2021, 2023); Nagel and Xu (2021); De la O, Han and Myers (2023)).

However, a different mechanism may underlie these stylized facts: prices may impact analyst cash flow expectations. For example, assume a decrease in discount rate raises stock price, while *investor* cash flow expectations do not change. *Analysts* do not know price rose because the discount rate fell, and instead raise their cash flow expectations so their personal valuations match the higher price. Thus, analyst expectations correlate with price. Since these expectations are "too high," they predict future forecast errors and returns. Yet these biased analyst cash flow expectations did not *cause* the price change; they simply *reflect* the price change caused by the discount rate decrease.

In this paper I present the first evidence that prices impact analyst cash flow expectations. Using two instruments based on Russell index reconstitution (Chang, Hong and Liskovich (2014); Pavlova and Sikorskaya (2023)) and mutual fund flow-induced trading (Lou (2012); Li (2021); Ben-David et al. (2022); Li, Fu and Chaudhary (2022)), I find in the cross section of equities that an exogenous 1% price increase raises analyst long-term earnings growth (LTG) expectations by 5 basis points and one to four year earnings-per-share (EPS) expectations and forecast errors by 20 to 40 basis points. These increases are persistent: analysts do not ex-post revise their expectations downward over the next year. These increases also appear permanent in the term structure of analyst expectations: they do not shrink as forecast horizon grows.

This impact of prices on analyst cash flow expectations is economically significant: it explains about half of the covariance between these objects. As Figure 1 displays, this mechanism explains 60% of the cross-sectional covariance of prices with LTG expectations, and 40% of the covariance with one to four year EPS expectations. Thus, this mechanism is quantitatively as important as the mechanism previous work focuses on (which explains the remainder of these covariances): that information and sentiment shocks to analyst expectations are shared by investors, and so impact prices. Since these results are in the cross

Figure 1: Decomposition of Covariance Between Analyst Cash Flow Expectations and Prices



This figure displays the proportion of the covariance of quarterly analyst cash flow expectation changes with contemporaneous price changes explained by the impact of prices on analyst cash flow expectations (red) and by common information or sentiment shocks to analyst and investor expectations (grey). These proportions sum to one. The left and right panels display this decomposition for changes in long-term earnings growth expectations and revisions to one to four year EPS expectations, respectively. Error bars represent quarterly block-bootstrapped 95% confidence intervals. See Section 5.3 for details.

section, there remains an open question of how prices impact cash flow expectations for the market.

These results are inconsistent with typical models used to interpret analyst cash flow expectations[1], which feature two core assumptions: 1) homogeneous analyst and investor beliefs, and 2) no feedback from prices to cash flow expectations (only "biased learning from exogenous fundamentals" in the terms of Bastianello and Fontanier (2021*b*)).

However, these results are consistent with two alternative model classes that relax each of these assumptions. The first class relaxes the assumption of homogeneous beliefs, such as models with dispersed information (e.g. Grossman and Stiglitz (1980); Hellwig (1980); Dubey, Geanakoplos and Shubik (1987); Kyle (1989); Jackson (1991); Mendel and Shleifer (2012)). In these models, at least some *investors'* cash flow expectations do not depend on prices, but *analysts* learn from prices, which reflect those investors' private information. The second class allows homogeneous cash flow expectations, but relaxes the assumption of no feedback from prices to these common expectations, such as models of price extrapolation (e.g. Jin and Sui (2022)). In these models, exogenous price rises today raise expectations of future prices, and agents mechanically raise *cash flow* expectations to justify these higher *price* expectations.

These latter two model classes present alternative interpretations of correlations of analyst cash flow

[1] La Porta (1996); De Bondt and Thaler (1990); Bordalo et al. (2019, 2022); De La O and Myers (2021, 2023); Nagel and Xu (2021); De la O, Han and Myers (2023)

expectations with prices, future returns, and forecast errors. In the models used in most previous work, biases in analyst cash flow expectations *drive* price variation. In these latter two classes, other variables (e.g. noise shocks, discount rates, biased price expectations) can drive prices, and analyst expectations may simply *reflect* price variation.

My results imply these alternative interpretations should be taken seriously when working with analyst cash flow expectations, and so motivate further study of heterogeneous beliefs and endogenous belief formation mechanisms.

The key challenge in measuring the impact of prices on analyst cash flow expectations is finding exogenous variation in prices. I need price variation due to noise trading unrelated to cash flow news, which creates omitted variable bias by impacting analyst expectations (directly) and prices (via investor expectations).

To tackle this challenge, I use two price instruments from previous work: benchmarking intensity changes around Russell index reconstitutions (Pavlova and Sikorskaya (2023)) and mutual fund flow-induced trading (Lou (2012)). These instruments use different assumptions, variation, and samples, yet yield similar estimates of the impact of prices on analyst cash flow expectations, which suggests this result is robust.

First, I use changes in benchmarking intensity around Russell index reconstitutions, following Pavlova and Sikorskaya (2023). Each June, Russell ranks stocks based on May-end market capitalization and assigns them to the Russell 1000 and 2000 indices based on a mechanical rank cutoff: those above the cutoff are assigned to the Russell 1000, and those below are assigned to the Russell 2000. Historically, more institutional capital has been benchmarked to the Russell 2000 than 1000. Hence, a Russell 1000 stock whose May market cap falls just below the cutoff moves to the Russell 2000, undergoes inflows of institutional capital due to benchmarking, and experiences positive returns in the reconstitution month of June (the "index effect"). Conditional on the May-end market cap, these reconstitution returns are exogenous to cash flow news (Chang, Hong and Liskovich (2014); Crane, Michenaud and Weston (2016); Glossner (2019)).

Pavlova and Sikorskaya (2023) note stocks that switch between the Russell 1000 and 2000 Blend indices also switch between the Russell 1000 and 2000 Value or Growth indices, which have different levels of benchmarked capital. Thus, two stocks moving from the Russell 1000 to 2000 can face different benchmarking flows, and so different reconstitution returns. To address this heterogeneity, Pavlova and Sikorskaya (2023) measure changes in benchmarking intensity (BMI) — a stock's total inelastic demand from all benchmarked managers — due to Russell reconstitution. Stocks with larger BMI increases face more price pressure.

Following Pavlova and Sikorskaya (2023), I use June BMI changes for stocks in a narrow window around the Russell market cap cutoffs to instrument for price. I find an exogenous, reconstitution-driven 1% price increase raises analyst one to four year EPS expectations by 40 basis points.

Second, I use the mutual fund flow-induced trading (FIT) instrument of Lou (2012) (similar to the flow-to-stock instrument of Wardlaw (2020)). Flows induce funds to do some mechanical rebalancing: funds tend to scale their preexisting holdings proportionally in response to flows. This predicted mechanical component of the cross-sectional trading induced by flows is uninformed and can provide exogenous price variation. An exogenous, flow-driven 1% price increase raises analyst one to four-year EPS expectations by 20 basis points, which is not statistically distinct from the 40 basis points estimate from the BMI instrument and does not shrink as forecast horizon grows. Moreover, the greater coverage of the FIT instrument (all stocks held by mutual funds) than the BMI instrument (stocks in narrow windows around Russell market cap cutoffs) also allows precise measurement of the impact of prices on LTG expectations, which far fewer analysts report than annual EPS expectations. An exogenous 1% price increase raises LTG expectations by 5 basis points. For both expectations types, these increases do not revert over the next year.

The FIT instrument does *not* require mutual fund flows to be exogenous. Flows are "aggregate shocks" within each quarter and do not create cross-sectional variation in the FIT instrument across stocks. Heterogeneous ownership shares create variation by providing heterogeneous exposures to these aggregate shocks across stocks: stocks owned in greater proportion by a fund are more exposed to its flows. Hence, the identifying assumption is ownership shares must not correlate with analyst cash flow expectations shocks across stocks. This condition is a special case of the result that exogenous shares are sufficient for a shift-share instrument to be exogenous (Goldsmith-Pinkham, Sorkin and Swift (2020)). Ownership shares and analyst belief shocks depending on common stock characteristics can violate this exogenous shares assumption. Controlling for characteristics (interacted with time fixed effects) associated with the ownership shares of funds that drive most of the FIT instrument variation yields similar results. Controlling for unobserved characteristics from a latent factor model that explain most ownership share variation also yields similar results. Funds holding portfolios of few stocks can raise similar issues, so I construct the FIT instrument from only funds with many holdings and find similar results. Systematic deviations from the proportional trading assumption can also raise similar issues (Berger (2023)), so I construct the instrument only from passive funds, which deviate far less from proportional trading than active funds, and find similar results.

This paper proceeds as follows. Section 2 introduces my test for if prices impact analyst cash flow expectations and explains how this impact alters interpretations of these expectations. Section 3 discusses the data. Sections 4 and 5 present evidence of price impact on analyst cash flow expectations using benchmarking intensity changes around Russell index reconstitutions and the FIT instrument. Section 6 concludes.

1.1 Related Literature

This paper relates to three bodies of literature: empirical work that uses analyst expectations as a proxy for investor beliefs, theoretical work on the impact of prices on cash flow expectations, and previous work that examines how different economic agents respond to exogenous price changes.

First, this paper relates to a large literature going back at least to Malkiel (1970) that uses analyst cash flow expectations as a proxy for investor expectations.[2] Most relevant is a growing body of work using analyst cash flow expectations to test asset pricing models in which investor beliefs deviate from rational expectations. A main conclusion of this literature is variation in investors' subjective cash flow expectations can explain much — if not all — of the time series and cross-sectional variation in stock prices. Chen, Da and Zhao (2013) find short- and long-term analyst earnings expectations revisions can explain most price variation in the cross section of stocks and aggregate market. De La O and Myers (2021) and De la O, Han and Myers (2023) find short-term analyst cash flow expectations explain much of the variation in aggregate-market and cross-sectional valuation ratios. De La O and Myers (2023) find forecast errors in short-term analyst cash flow expectations and long-term professional forecaster inflation expectations explain aggregate stock and bond market price variation. La Porta (1996), Bordalo et al. (2019), and Bordalo et al. (2022) find analyst LTG expectations correlate with valuation ratios in the cross section of stocks and negatively predict future returns. Nagel and Xu (2021) and Bordalo et al. (2022) find similar results for the aggregate market. De Bondt and Thaler (1990), La Porta (1996), Bordalo et al. (2019), and Bordalo et al. (2022) find predictable analyst expectation revisions and errors for the cross section of stocks and aggregate market.

These results contradict rational expectations models, in which discount rate variation explains most aggregate market price variation and expectation revisions and errors are not predictable. Thus, these results motivate models in which investors share analyst cash flow expectations that deviate from rational expectations and distort prices (Bordalo et al. (2019, 2022); Nagel and Xu (2021); De La O and Myers (2021, 2023); De la O, Han and Myers (2023)). In these models, prices do not impact cash flow expectations; biases arise only from biased learning from exogenous fundamentals (Bastianello and Fontanier (2021*b*)).

This paper provides evidence of a different mechanism underlying correlations of analyst cash flow

[2] E.g. Frankel and Lee (1998); Lee, Myers and Swaminathan (1999); Lee and Swaminathan (2000); Hribar and McInnis (2012); Bouchaud et al. (2019); Brandon and Wang (2020); Landier and Thesmar (2020) among many others.

expectations with prices, future returns, and forecast errors: prices impact analyst cash flow expectations. While much of the literature overlooks this mechanism, the contribution of this paper is to provide the first evidence that prices impact analyst cash flow expectations and demonstrate this mechanism is quantitatively important.[3] These results are inconsistent with the aforementioned models with homogeneous analyst and investor cash flow expectations that do not depend on prices. Moreover, these results suggest biases in analyst cash flow expectations may not distort prices, but may simply reflect distortions from other variables (e.g. noise shocks, discount rates, biased price expectations). Since my results are in the cross section, they cannot speak directly to time series correlations of analyst expectations with the aggregate market.

Second, this paper relates to theoretical work on the impact of prices on cash flow expectations. The price impact on analyst cash flow expectations I document is consistent with models featuring learning from prices (e.g. Grossman and Stiglitz (1980); Hellwig (1980); Dubey, Geanakoplos and Shubik (1987); Kyle (1989); Jackson (1991); Mendel and Shleifer (2012)).[4] However, previous work using analyst expectations as a proxy for investor beliefs has generally not accounted for this learning from prices mechanism. This paper provides the first evidence that analyst cash flow expectations behave in a manner consistent with learning from prices. This price impact on analyst cash flow expectations is also consistent with certain models of price extrapolation (e.g. Jin and Sui (2022)).

Third, this paper relates to work on how different agents respond to exogenous price changes. Previous work uses mutual fund flow-driven price pressure to examine the impact of prices on various corporate finance outcomes.[5] A separate literature uses Russell index reconstitutions to study the impact of institutional and passive ownership on corporate governance and product market outcomes.[6] This paper uses both instruments to study how exogenous price changes impact analyst cash flow expectations. Importantly, I use the Lou (2012) mutual fund flow-induced trading instrument for prices, which is not subject to the Wardlaw (2020) critique of the Edmans, Goldstein and Jiang (2012) version of this instrument.[7]

[3]Some instances of previous work consider this mechanism. E.g. Malkiel (1970) muses: "The strong correlation between price-earnings multiples and predicted growth rates leads one to question the line of causality. Do stocks with high expected growth rates tend to sell at high price- earnings multiples because investors actively bid up the shares of companies with favorable prospects? Or does the security analyst see a large price-earnings ratio in the market and decide from this that the firm in question must indeed be a 'growth stock?'"

[4]Glaeser and Nathanson (2017); Bastianello and Fontanier (2021*b*,*a*); Bordalo et al. (2021) study mislearning from prices.

[5]E.g. Seasoned equity issuance (Giammarino et al. (2004); Khan, Kogan and Serafeim (2012)), M&A (Edmans, Goldstein and Jiang (2012); Eckbo, Makaew and Thorburn (2018)), payout policy (Derrien, Kecskés and Thesmar (2013)), R&D spending (Phillips and Zhdanov (2013)), shareholder activism (Norli, Ostergaard and Schindele (2015)), management earnings forecasts (Zuo (2016)), analyst coverage (Lee and So (2017)), and investment Lou and Wang (2018); Dessaint et al. (2019).

[6]E.g. Schmidt and Fahlenbrach (2017); Appel, Gormley and Keim (2016, 2019, 2021); Heath et al. (2022); Sharma (2023)

[7]Wardlaw (2020) demonstrates the Edmans, Goldstein and Jiang (2012) construction of mutual fund flow-induced trading mechanically depends on the current-period return and argues this dependence threatens the instrument's exogeneity.

Figure 2: Within Stock-Quarter Timeline

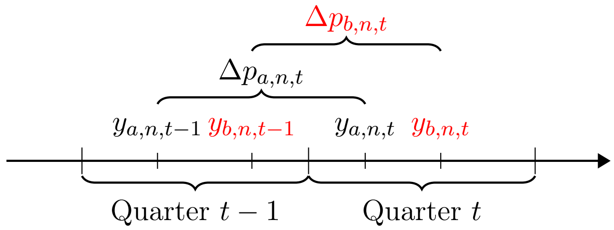


Illustration of staggered timing of expectation releases for two analysts a and b for the same stock n.

2 Identifying Price Impact on Analyst Expectations and Why it Matters

This section introduces my test for if prices impact analyst cash flow expectations. I first explain the challenge in measuring this impact and why exogenous price variation proves necessary. I then present three models of analyst expectations to explain how this impact alters interpretations of these expectations.

2.1 The Test

The goal of this paper is to determine if prices impact analyst cash flow expectations. The challenge in measuring this impact is common information or sentiment shocks create omitted variable bias by impacting analyst expectations (directly) and prices (via investor expectations). Consider this system of equations:

$$\begin{aligned} \Delta p_{a,n,t} &= M z_{a,n,t} + \epsilon_{a,n,t} \\ \Delta y_{a,n,t} &= \alpha \Delta p_{a,n,t} + \nu_{a,n,t} \end{aligned} \tag{1}$$

$\Delta y_{a,n,t}$ is the quarterly change in analyst a's expectations for stock n in quarter t. $\Delta p_{a,n,t}$ is the contemporaneous percentage price change (ex-dividend return) between the two quarterly report dates for analyst a for stock n in quarters $t-1$ and t (the a subscript indicates different analysts report expectations for stock n in quarter t on different days, and so face different inter-announcement price changes, as in Figure 2).

α is the impact of prices on analyst cash flow expectations. I want to test if α is positive. A regression of analyst expectations on prices does not identify α because both variables experience other correlated shocks $\epsilon_{a,n,t}$ and $\nu_{a,n,t}$, which create omitted variable bias. For example, public signals about cash flows (e.g. EPS announcements) that both investors and analysts learn from would appear in $\epsilon_{a,n,t}$ and $\nu_{a,n,t}$.

Yet if a noise-trader demand shock $z_{a,n,t}$ provides exogenous price variation uncorrelated with analyst cash flow expectations shocks $\nu_{a,n,t}$, then the following two-stage least squares regression identifies α:

$$\Delta p_{a,n,t} = a_1 z_{a,n,t} + e_{1,a,n,t}$$

$$\Delta y_{a,n,t} = \alpha \Delta \hat{p}_{a,n,t} + e_{2,a,n,t}$$

To obtain a consistent estimate of α, the price instrument $z_{a,n,t}$ must satisfy:

1. (Relevance) $M \neq 0$ in (1): The instrument has an effect on price.

2. (Exogeneity) $\mathbb{E}[z_{a,n,t}\nu_{a,n,t}] = 0$: The instrument affects analyst cash flow expectations only through price; it does not correlate with other shocks to analyst expectations. For example, if $z_{a,n,t}$ correlates with public information (e.g. EPS announcements) that both investors and analysts learn from, this condition is violated. Thus, $z_{a,n,t}$ must provide variation in prices that is unrelated to cash flow news.

To gauge the economic significance of α, I calculate the proportion of the cross-sectional covariance of prices with analyst cash flow expectations due to the impact of prices on analyst expectations:

$$\frac{\alpha \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y_{a,n,t}\right)} = \frac{\text{Two Stage Least Squares Estimate of } \alpha}{\text{OLS Coefficient in Regression of } \Delta y_{a,n,t} \text{ on } \Delta p_{a,n,t}}. \tag{2}$$

V^{CX} and Cov^{CX} are the cross-sectional variance and covariance after removing time fixed effects.

Sections 4 and 5 describe strategies to identify α using different empirical noise trader demand shocks.

2.2 Three Models to Interpret Analyst Cash Flow Expectations

I present three stylized models that match common facts about analyst cash-flow expectations: they correlate positively with prices[8], negatively predict future returns[9], and feature predictable forecast errors.[10]

The first model features the two core assumptions previous work[11] uses to interpret these facts: homogeneous analyst and investor cash flow expectations and no feedback from prices to these expectations. The second model relaxes the first assumption of belief homogeneity and involves analyst — but not investor — learning from prices due to dispersed information. The third model relaxes the second assumption of no

[8] Chen, Da and Zhao (2013); De La O and Myers (2021); Bordalo et al. (2022); De la O, Han and Myers (2023)
[9] La Porta (1996); Bordalo et al. (2019, 2022)
[10] De Bondt and Thaler (1990); La Porta (1996); Bordalo et al. (2019, 2022)
[11] Bordalo et al. (2019, 2022); Nagel and Xu (2021); De La O and Myers (2021, 2023); De la O, Han and Myers (2023).

feedback from prices and involves homogeneous investor and analyst cash flow expectations that depend on prices due to price extrapolation.

In the first model, prices do not impact analyst cash flow expectations, which are shared by investors and so distort prices. In the latter two models, prices impact analyst expectations. Biases in these expectations do not distort prices, but *reflect* distortions from other variables (e.g. discount rates).

Thus, if prices impact analyst cash flow expectations, it is possible biases in these expectations reflect price variation instead of driving it. We must consider heterogeneous beliefs (as in the second model) and endogenous belief formation mechanisms (as in the third model) when interpreting these expectations.

2.2.1 Homogeneous Cash Flow Expectations and No Feedback from Prices

This model features the two assumptions previous work uses to interpret analyst expectations: homogeneous analyst and investor beliefs and no feedback from prices to these beliefs. See Appendix A.1 for proofs.

There are two periods. There is one asset with fixed supply of one share and random period-two payoff:

$$\tilde{D}_2 = \bar{D} + \tilde{\varepsilon}, \tilde{\varepsilon} \sim N(0, \sigma_\varepsilon^2).$$

All tilde quantities represent random variables. I normalize the risk-free rate to zero.

A representative investor has initial wealth W_1 and constant absolute risk aversion (CARA) utility over terminal wealth $\tilde{W}_2$:

$$U(\tilde{W}_2) = -\exp\left[-A\left(\tilde{W}_2\right)\right] = -\exp\left[-A\left(W_1 + Q_1^I\left(\tilde{D}_2 - \tilde{P}_1\right)\right)\right].$$

Q_1^I is the number of shares demanded in period one. $\tilde{P}_1$ is the period-one price.

The investor and an analyst have the same biased expectation:

$$\mathbb{E}_1^I\left[\tilde{D}_2\right] = \mathbb{E}_1^A\left[\tilde{D}_2\right] = \bar{D} + \tilde{s}, \tilde{s} \sim N\left(0, \sigma_s^2\right), \tag{3}$$

$\tilde{s}$ is a random sentiment shock in period one that does not depend on price.[12] The market clearing price is:

$$\tilde{P}_1 = \bar{D} + \tilde{s} - A\sigma_\varepsilon^2. \tag{4}$$

[12]This sentiment shock serves the same purpose as an exogenous signal of $\tilde{D}_2$ that the investor overweights (here $\tilde{s}$ doesn't predict $\tilde{D}_2$, so the optimal weight is zero).

This model matches the core set of stylized facts about analyst expectations.

Proposition 1. *Prices correlate positively with analyst cash flow expectations:* $Cov\left(\tilde{P}_1, \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) > 0$. *The period-one analyst expectation and price negatively predict the period-two return:* $Cov\left(\mathbb{E}^A\left[\tilde{D}_2\right], \tilde{D}_2 - \tilde{P}_1\right) = Cov\left(\tilde{P}_1, \tilde{D}_2 - \tilde{P}_1\right) < 0$. *The period-one analyst expectation and price negatively predict the period-two forecast error:* $Cov\left(\mathbb{E}_1^A\left[\tilde{D}_2\right], \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) = Cov\left(\tilde{P}_1, \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) < 0$.

In this model, analysts and investors have the same biased cash flow expectations that distort prices but do not depend on prices.

2.2.2 Relaxing Homogeneous Expectations: Dispersed Information

Consider an alternative model in which analysts learn from prices because investors have different beliefs. See Appendix A.2 for proofs.

I make four changes to the previous model. First, I add a new shock $\tilde{\eta}$ to the random period-two payoff:

$$\tilde{D}_2 = \bar{D} + \tilde{\eta} + \tilde{\varepsilon}, \tilde{\eta} \sim N\left(0, \tau^{-1}\right), \tilde{\varepsilon} \sim N\left(0, \sigma_\varepsilon^2\right).$$

Second, the CARA utility investor has rational expectations and knows $\tilde{\eta}$:

$$\mathbb{E}_1^I\left[\tilde{D}_2\right] = \bar{D} + \tilde{\eta}.$$

Third, I introduce a noise trader with demand: $Q_1^N = \tilde{z} \sim N\left(0, \sigma_z^2\right)$. The market price is:

$$\tilde{P}_1 = \bar{D} + \tilde{\eta} - (1 - \tilde{z})\, A\sigma_\varepsilon^2. \tag{5}$$

Fourth, there is an analyst who reports an expectation for $\tilde{D}_2$ at the end of period one after observing the price. The analyst does *not* know $\tilde{\eta}$, the private information of the CARA investor, nor the quantity demanded by either investor. Given (5), price is a noisy signal of $\tilde{\eta}$ and the analyst will learn from the price.

The analyst uses Bayes rule, but has biased beliefs about the noise trader demand shock. The analyst underestimates the noise trader demand shock variance, believing the noise shock is distributed as $\tilde{z} \sim N\left(0, \hat{\sigma}_z^2\right)$, where $\hat{\sigma}_z < \sigma_z$. Thus, he overestimates how informative price is about the CARA investor's private information, and so overreacts to price. Without this assumption, analyst forecast errors are not predictable, but analyst expectations still correlate with prices and predict future returns.

The analyst has the objectively correct prior for $\tilde{\eta}$, believing $\tilde{\eta} \sim N(0, \tau^{-1})$. Thus, letting $\hat{\omega}^{-1} = A^2\sigma_\varepsilon^4\hat{\sigma}_z^2$, Bayes rule yields the following analyst posterior expectation:

$$\mathbb{E}_1^A\left[\tilde{D}_2\right] = \bar{D} + \frac{\hat{\omega}}{\tau + \hat{\omega}}\left(\tilde{P}_1 - \bar{D} + A\sigma_\varepsilon^2\right). \tag{6}$$

This model matches the same stylized facts about analyst expectations as the model in Section 2.2.1.

Proposition 2. *Prices correlate positively with analyst cash flow expectations: Cov* $\left(\tilde{P}_1, \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) > 0$. *The period-one analyst expectation and price negatively predict the period-two return: Cov* $\left(\mathbb{E}^A\left[\tilde{D}_2\right], \tilde{D}_2 - \tilde{P}_1\right)$, *Cov* $\left(P_1, \tilde{D}_2 - \tilde{P}_1\right) < 0$. *The period-one analyst expectation and price negatively predict the period-two forecast error: Cov* $\left(\mathbb{E}_1^A\left[\tilde{D}_2\right], \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right)$, *Cov* $\left(\tilde{P}_1, \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) < 0$.

In this model, biases in analyst expectations do not distort price; they simply reflect distortions caused by the noise shock. Investors and analysts have different expectations; the noise shock $\tilde{z}$ drives a wedge between these agents' expectations. Thus, the correlations of analyst expectations with prices, future returns, and analyst forecast errors do not shed light on investor rationality or the determinants of asset prices.

Shocks to risk aversion (A) or volatility (σ_ε) that the analyst does not observe would serve the same purpose as the noise trader demand shock. Thus, analyst learning from price implies that expected return variation (i.e. discount rate variation) can drive a wedge between analyst and investor cash flow expectations.

For simplicity, I assume the analyst has no private information and so his expectation does not impact price.[13] One could easily extend this model to include analyst private information, in which case the analyst expectation would impact price because the CARA investor would rationally learn from the analyst report.

In a more sophisticated model with heterogeneous investors, analyst expectations may align with those of "uninformed" investors who learn from prices, but not with those of "informed" investors who do not.

2.2.3 Relaxing no Feedback from Prices: Price Extrapolation

Now consider a third model in which analysts and investors have the same cash flow expectations, which depend on price due to price extrapolation (as in Jin and Sui (2022)). See Appendix A.3 for proofs.

[13] A large literature documents high-frequency price responses to analyst report releases (Davies and Canes (1978); Groth et al. (1979); Barber and Loeffler (1993); Stickel (1995); Albert Jr and Smaby (1996); Francis and Soffer (1997); Park and Stice (2000); Barber et al. (2001); Brav and Lehavy (2003); Irvine (2003); Asquith, Mikhail and Au (2005); Kerl and Walter (2008); Fang and Yasuda (2014); Ishigami and Takeda (2018)).

I make four changes to the Section 2.2.1 model. First, there are three periods and one cash flow in $t = 3$:

$$\tilde{D}_3 = \bar{D} + \tilde{\varepsilon}, \tilde{\varepsilon} \sim N\left(0, \sigma_\varepsilon^2\right).$$

Second, there is a noise trader with demand in periods $t = 1$ and 2 of: $Q_t^N = \tilde{z}_t \sim N\left(0, \sigma_z^2\right)$.

Third, the CARA investor and analyst have rational expectations *in period two* about $\tilde{D}_3$:

$$\mathbb{E}_2^I\left[\tilde{D}_3\right] = \mathbb{E}_2^A\left[\tilde{D}_3\right] = \bar{D}.$$

However, both agents have extrapolative expectations about the period-two price in period one:

$$\mathbb{E}_1^I\left[\tilde{P}_2\right] = \mathbb{E}_1^A\left[\tilde{P}_2\right] = \mathbb{E}_1^{Obj}\left[\tilde{P}_2\right] + \phi\tilde{P}_1, \tag{7}$$

where $\mathbb{E}_1^{Obj}$ represents the objective expectation under the true data generating process and $\phi > 0$.

The market clearing prices in each period are:

$$\tilde{P}_2 = \bar{D} - \left(1 - \tilde{z}_2\right) A\sigma_\varepsilon^2 \tag{8}$$

$$\tilde{P}_1 = \frac{1}{1-\phi}\left[\bar{D} - A\sigma_\varepsilon^2 - \left(1 - \tilde{z}_1\right) A^3\sigma_\varepsilon^4\sigma_z^2\right]. \tag{9}$$

Fourth, both agents form *period-one* cash flow expectations to rationalize the biased price expectations:

$$\mathbb{E}_1^I\left[\tilde{D}_3\right] = \mathbb{E}_1^A\left[\tilde{D}_3\right] = \bar{D} + \phi\tilde{P}_1 \tag{10}$$

Note that without extrapolation ($\phi = 0$), this investor has rational expectations about $\tilde{D}_3$ in period one. However, with extrapolation ($\phi > 0$), the period-one noise shock biases cash flow expectations.

This model matches the same stylized facts about analyst expectations as the model in Section 2.2.1.

Proposition 3. *Prices correlate positively with analyst cash flow expectations: $Cov\left(\tilde{P}_1, \mathbb{E}_1^A\left[\tilde{D}_3\right]\right) > 0$. The period-one analyst expectation and price negatively predict the period-two return: $Cov\left(\mathbb{E}_1^A\left[\tilde{D}_3\right], \tilde{P}_2 - \tilde{P}_1\right)$, $Cov\left(\tilde{P}_1, \tilde{P}_2 - \tilde{P}_1\right) < 0$. The period-one analyst expectation and price negatively predict the period-three forecast error: $Cov\left(\mathbb{E}_1^A\left[\tilde{D}_3\right], \tilde{D}_3 - \mathbb{E}_1^A\left[\tilde{D}_3\right]\right)$, $Cov\left(\tilde{P}_1, \tilde{D}_3 - \mathbb{E}_1^A\left[\tilde{D}_3\right]\right) < 0$.*

In this model, analysts and investors have the same biased, period-one cash flow expectations. Yet these biases do not distort price; they reflect distortions due to the noise shock and biased *price* expectations.

3 Data

This paper uses four main data sources: analyst cash flow expectations, stock prices, Russell index constituents, and mutual fund holdings and flows.

I use two sets of analyst cash flow expectations from I/B/E/S. First, I use the long-term earnings-per-share (EPS) growth (LTG) expectations focused on by Bordalo et al. (2019), Nagel and Xu (2021), and Bordalo et al. (2022). I/B/E/S defines the LTG expectations as representing analysts' expected annual EPS growth over a firm's "next full business cycle," which I/B/E/S describes as three to five years (Wharton Research Data Services (2008)), though some work argues these expectations capture growth expectations over longer horizons of up to five to ten years (Sharpe (2005)). I/B/E/S reports LTG expectations at the stock $\times$ analyst institution $\times$ analyst $\times$ quarter level. I average LTG expectations for each stock within each quarter at the analyst institution level and take quarterly differences to obtain a stock $\times$ analyst institution $\times$ quarter panel of quarterly changes in LTG expectations.[14]

Second, I use the annual EPS expectations over shorter horizons of one to four years focused on by De La O and Myers (2021, 2023), and De la O, Han and Myers (2023). I/B/E/S reports EPS forecasts at the stock $\times$ fiscal year horizon $\times$ analyst institution $\times$ analyst $\times$ quarter level. For example, in the second quarter of 2022 I see the Apple annual EPS forecasts issued by all equity research analysts at Goldman Sachs for fiscal years 2022, 2023, etc. Forecast horizons extend up to ten fiscal years ahead, but coverage declines with horizon (sharply after two years), and so I focus on the one to four year horizons. For each horizon, I average EPS forecasts for each stock within each quarter at the analyst institution level (e.g. I average the EPS forecasts for fiscal year 2022 for Apple made by all Goldman Sachs analysts during the second quarter of 2022). I then linearly interpolate among horizons to construct fixed h-year horizon EPS forecasts. For example, to obtain the one-year EPS forecast from June 2022 to June 2023, I interpolate between the fiscal year 2022 and 2023 EPS forecasts.[15] I then construct quarter-over-quarter EPS expectation revisions. Let $E_{a,n,t+4h|t}$ be the h-year ahead annual EPS expectation reported by analyst institution a for stock n in

[14]Using institution-level instead of analyst-level variation creates more quarter over quarter matches when computing quarterly expectations changes. I winsorize these final values at the 5% level (within each horizon h) to remove some extremely large outliers. Usually only one analyst at an institution covers stock n in quarter t. If multiple analysts in institution a cover stock n in quarter t, they usually report expectations on the same day. If multiple analysts from institution a report expectations for stock n in quarters $t-1$ or t on different days, I compute the quarterly inter-announcement price change ($\Delta p_{a,n,t}$) between the day of the first such expectation release in quarter $t-1$ and the day of the last such expectation release in quarter t.

[15]De La O and Myers (2021) follow the same interpolation procedure.

quarter t. I define the h-year ahead EPS expectation revision from quarter t to $t+1$ as:

$$\Delta E_{a,n,t+4h|t+1} \equiv \frac{E_{a,n,t+4h|t+1} - E_{a,n,t+4h|t}}{E_{a,n,t+4h|t}}. \tag{11}$$

I drop all observations where $E_{a,n,t+4h|t} \leq 0$. Thus, I obtain a stock $\times$ analyst institution $\times$ quarter panel of quarterly revisions in h-year EPS expectations, where h ranges from one to four years.[16]

The LTG and annual EPS expectations each have their own advantages. On one hand, present value identities suggest the LTG expectations should correlate more strongly with prices, as documented in previous work (Bordalo et al. (2019, 2022, 2021)). On the other hand, the annual expectations have far greater coverage, which enables more powerful tests. Moreover, the fixed forecast horizon means $\Delta E_{a,n,t+4h|t+1}$ directly captures forecast error changes:

$$\text{Forecast Error}_{a,n,t+4h|t} = \text{Realized EPS}_{n,t+4h} - E_{a,n,t+4h|t}$$
$$\text{Forecast Error}_{a,n,t+4h|t+1} - \text{Forecast Error}_{a,n,t+4h|t} = E_{a,n,t+4h|t} - E_{a,n,t+4h|t+1}.$$

The forecast error change is the negative of the expectation revision numerator in (11). Thus, the annual expectation revisions demonstrate that not only do analyst cash flow expectations rise in response to exogenous price increases, analyst *forecast errors* also rise (in magnitude, i.e. become more negative). Thus, analysts do not raise cash flow expectations in response to exogenous price increases solely because these increases raise actual future earnings.

I obtain stock price data from CRSP and accounting data to construct firm characteristics from the Compustat North America Fundamentals Annual and Quarterly Databases.

The authors of Pavlova and Sikorskaya (2023) provide benchmarking intensity and constituent data.

To construct the flow-induced trading instrument of Lou (2012), I use mutual fund holdings from the Thomson Reuters S12 database and mutual fund flows from the CRSP Mutual Fund database.[17]

4 Evidence from Index Reconstitutions

This section provides evidence that prices impact analyst cash flow expectations using my first exogenous price shock: changes in benchmarking intensity around annual June Russell index reconstitutions.

[16] I winsorize these final values at the 5% level to remove some extremely large outliers.
[17] Following Wardlaw (2020), I drop sector mutual funds when constructing the flow-induced trading instrument.

On a specified day in May, Russell ranks all eligible stocks by market capitalization. Stocks above a specific rank cutoff are assigned to the Russell 1000, and those below are assigned to the Russell 2000. Historically, more institutional capital has been benchmarked to the Russell 2000 than 1000. Thus, a stock from the Russell 1000 in year $t-1$ whose market cap falls just below the cutoff in year t will move to the Russell 2000 in June, undergo inflows of institutional capital due to benchmarking, and experience positive returns in June. Similarly, a stock from the Russell 2000 in year $t-1$ whose market cap falls just above the cutoff will move to the Russell 1000 and experience outflows and negative returns. Conditional on the May rank-date market cap, Russell index membership in June is exogenous to fundamental news (Chang, Hong and Liskovich (2014); Crane, Michenaud and Weston (2016); Glossner (2019)). Thus, the June returns induced by this index reconstitution are also exogenous to fundamental news.

Pavlova and Sikorskaya (2023) note these reconstitution returns differ across stocks. Every stock in the Russell 2000 Blend index is also in the Russell 2000 Value or Growth indices, which have different levels of benchmarked capital. Every stock in the Russell 1000 Blend index is also in the Russell 1000 Value or Growth indices, and some (those under market cap rank 200) are in the Russell Midcap Blend, Value, and Growth indices. Thus, a stock moving from the Russell 1000 Value to the Russell 2000 Value may experience different inflows of benchmarked capital — and so different price pressure — than a stock moving from the Russell 1000 Growth to the Russell 2000 Growth.

The Pavlova and Sikorskaya (2023) benchmarking intensity (BMI) measure captures this heterogeneity:

$$BMI_{n,t} = \sum_{\text{Index } j} \frac{\text{Institutional AUM Benchmarked to Index } j_t \cdot \mathbf{1}\left(n \in \text{Index } j_t\right)}{\text{Index } j \text{ Market Value}_t}.$$

$BMI_{n,t}$ captures the inelastic demand for stock n in month t by all benchmarked mutual funds and ETFs. It depends on which indices j the stock is part of and the proportion of the total market value of index j that is held by benchmarked investors. Pavlova and Sikorskaya (2023) construct BMI from thirty-four indices that account for about 90% of mutual fund and ETF assets, including the nine Russell benchmarks.

I use June BMI changes in each year for stocks in a narrow window around the Russell 1000/2000 reconstitution thresholds as a price instrument. Stocks with larger BMI changes experience more benchmarking inflows and more price pressure. While BMI is generally endogenous because index membership is, June BMI *changes* for stocks in this window are driven by Russell index membership changes, which are exogenous to fundamental news conditional on the the May rank-date market cap. Thus, $\Delta BMI_{n,t}$ satisfies the exogeneity condition and is uncorrelated with analyst belief shocks (conditional on stock-level controls $\boldsymbol{X}_{n,t}$ discussed below):

$$\mathbb{E}\left[\Delta BMI_{n,t}\nu_{a,n,t} \mid \boldsymbol{X}_{n,t}\right] = 0, \forall a, t.$$

Note the importance of controlling for the rank-date market cap. Bad news before the rank date could impact analyst beliefs (directly) and $\Delta BMI_{n,t}$ (by lowering the market cap and moving stock n from the Russell 1000 to 2000). This situation threatens exogeneity. Yet since the bad news only impacts $\Delta BMI_{n,t}$ through the rank-date market cap, controlling for that market cap makes $\Delta BMI_{n,t}$ conditionally exogenous.

One may be concerned that BMI increases raise cash flow expectations directly (instead of through prices) if analysts pay attention to passive ownership increases and expect them to improve corporate governance or product market outcomes. These scenarios would threaten exogeneity. However, previous work finds mixed results for the effect of passive ownership increases on corporate governance quality.[18] Moreover, switching from the Russell 1000 to 2000 correlates *negatively* or not at all with future profitability and cash flows (Pavlova and Sikorskaya (2023); Sharma (2023)).[19] Hence, it is unlikely analysts *raise* their cash flow expectations directly due to the passive ownership increases accompanying BMI increases. Section 4.3 conducts additional tests to address this concern.

I run the following two-stage least squares regression:

$$\begin{aligned}
\Delta p_{a,n,t} &= a_1 \Delta BMI_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t} \\
\Delta E_{a,n,t+4h|t} &= \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{\beta}_2^{'} \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t,h}.
\end{aligned} \tag{12}$$

The first stage regresses quarterly percentage price changes between analyst reports on June BMI changes.[20] The second stage regresses quarterly revisions to annual EPS expectations on instrumented price changes.

I restrict the sample to analyst expectations changes exposed to the June index reconstitutions: those for which the original expectation is reported in June or earlier and the revised expectation is reported in June or later, as in Figure 3. I include all annual EPS expectations revisions with horizons up to four years.

I also restrict the sample to stocks within a narrow bandwidth around the reconstitution market cap cutoffs. I use 150 stocks around each cutoff in the baseline analysis (Section 4.3 finds similar results for alternative bandwidths). Prior to 2007, the rank cutoff was the 1,000th stock. To reduce turnover, since 2007 Russell has used a "banding policy" under which there are two separate cutoffs for stocks starting

[18] E.g. Schmidt and Fahlenbrach (2017); Appel, Gormley and Keim (2016, 2019, 2021); Heath et al. (2022)

[19] Sharma (2023) finds switching from the Russell 1000 to 2000 is weakly associated with *lower* profitability and cash flows over the next year. Pavlova and Sikorskaya (2023) finds "little evidence" that ΔBMI correlates with future cash flow changes.

[20] Following Pavlova and Sikorskaya (2023), I winsorize price changes at 1%. Section 4.3 finds similar results at 0%.

Figure 3: Russell Reconstitution Timeline

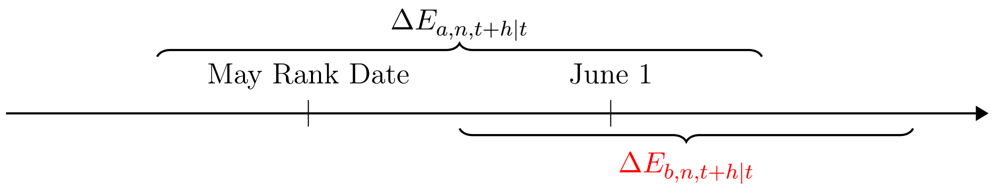


Illustration of Russell index reconstitution timing. $\Delta E_{a,n,t+h|t}$ is an h-year EPS expectation change for analyst a and stock n, where the original expectation was reported prior to the May rank date and the revised expectation is reported after June 1st. $\Delta E_{b,n,t+h|t}$ is an h-year expectation change for analyst b and stock n, where the original expectation was reported after the May rank date but before the end of June and the revised expectation is reported after June 1st. $\Delta E_{a,n,t+h|t}$ is included only in the full sample, while $\Delta E_{b,n,t+h|t}$ is included in both the full- and post-rank samples (discussed in Section 4.3).

in the Russell 1000 and 2000 pre-reconstitution, both of which are mechanical functions of the firm size distribution. Thus, there is a "band" of market caps including stocks from the Russell 1000 and 2000. Appendix B.1 explains the Russell methodology I use to calculate these cutoffs. Since Russell ranks stocks using a proprietary market cap that I lack access to, I use the method of Ben-David, Franzoni and Moussawi (2019) to approximate this proprietary market cap using standard databases.[21] Doing so predicts assignment to the Russell 1000 and 2000 with high accuracy, as shown in Appendix Table B1. Following previous work, I use May — not June — market caps to calculate the Russell reconstitution thresholds.[22]

$\boldsymbol{X}_{n,t}$ includes stock-level controls used by Pavlova and Sikorskaya (2023): May rank-date log market cap, one-year monthly average bid-ask percentage spread[23], and the banding controls from Appel, Gormley and Keim (2019) (an indicator for having rank-date market cap in the "band", an indicator for being in the Russell 2000 in May, and the interaction of these indicators). Whereas Pavlova and Sikorskaya (2023) use the proprietary Russell market cap, I calculate market cap from standard databases. These variables determine Russell 1000/2000 membership. Conditional on these controls, $\Delta BMI_{n,t}$ in June is exogenous. FE_t are year fixed effects.

In this restricted sample, there is enough power to quantify the impact of prices on annual EPS expectations, but not on LTG expectations, which have far less analyst coverage (see Appendix Table B3 for

[21] See Appendix Table B1 for details.

[22] E.g. Chang, Hong and Liskovich (2014); Appel, Gormley and Keim (2021); Wei and Young (2021))

[23] Pavlova and Sikorskaya (2023) note changes in a stock's liquidity can impact both its returns (by altering the liquidity premium) and BMI. Thus, they control for Russell's proprietary float factor and the rolling average bid-ask percentage spread (to address staleness in the float factor). Lacking access to Russell's proprietary float factor, I control for the bid-ask spread.

Table 1: Summary Statistics for Russell Reconstitution Instrument

	$\Delta BMI_{n.t}$	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	Num Stocks/Year	Percent Covered
Num Obs.	164,512.00	164,512.00	164,512.00	20.00	20.00
Mean	0.00	-0.02	0.02	432.45	0.89
Std. Dev	0.04	0.20	0.24	148.33	0.04
Min	-0.41	-0.63	-0.60	240.00	0.80
25%	-0.01	-0.09	-0.12	260.00	0.87
50%	0.00	0.00	0.00	541.50	0.91
75%	0.01	0.07	0.13	554.00	0.92
Max	0.29	0.51	0.90	562.00	0.94

Summary statistics for observations in the 150-stock window around Russell reconstitution market cap thresholds in each year for May-to-June changes in benchmarking intensity ($\Delta BMI_{n,t}$), quarter-over-quarter revisions in analyst EPS expectations for forecast horizons of one to four years ($\Delta E_{a,n,t+4h|t}$), inter-announcement percentage price changes between expectation releases in consecutive quarters ($\Delta p_{a,n,t}$), the number of stocks in the window in each year, and the percentage of all stocks in each window that I observe analyst expectations for. Expectations revisions and price changes are expressed in absolute terms (i.e. 0.01 is 1%). The time period is 1999-05:2018-09.

details).[24] I quantify the impact of prices on LTG expectations in Section 5 using mutual fund flow-induced trading to instrument for prices.

Table 1 presents summary statistics. There are $164,512$ total analyst-stock-horizon-year observations. The time period is 1999 to 2018, as this is the period in which I observe Russell index constituents in May (pre-reconstitution) and June (post-reconstitution). In the average year, I observe analyst expectation changes for about 90% of the firms in the 150-stock bandwidth around the reconstitution cutoffs.[25]

4.1 Empirical Results

Table 2 displays the baseline results. Column 1 displays the OLS regression of annual EPS expectation revisions on contemporaneous price changes, and finds a strong association between these objects, as documented in previous work (De La O and Myers (2021, 2023)). The first stage regression in column 2 is strong: Russell reconstitution-driven BMI increases raise prices. The partial F-statistic (11.7) is above the conventional threshold of 10. The reduced-form coefficient in column 3 is also significant: Russell reconstitution-driven BMI increases raise annual EPS expectation revisions. The second-stage α estimate in column 4 reveals a statistically and economically significant effect of prices on annual EPS expectation revisions: an

[24] In this restricted sample there are only $3,758$ analyst-stock-year observations for the LTG expectation changes, versus $164,512$ observations for the annual EPS expectation revisions. As Table B3 displays, the two-stage least squares estimate of α in this LTG expectation sample is $\alpha = 2.1$ basis points, which is economically significant and in the 95% confidence interval of the $\alpha = 5.5$ basis points estimate obtained from the flow-induced trading instrument (see Section 5.1 for details). However, the 95% confidence interval for this $\alpha = 2.1$ basis points estimate is wide (-22.7 to 26.9 basis points) due to the small sample.

[25] Since there is only one market cap cutoff before 2007, there are 300 stocks in the 150-stock window. After the introduction of the banding policy in 2007, there are two cutoffs, and so 600 stocks in the 150-stock window. See Appendix B.1 for details.

Table 2: Effect of Prices on Annual EPS Expectations Using ΔBMI as Instrument

	(1) OLS	(2) First Stage	(3) Reduced Form	(4) 2SLS
$\Delta p_{a,n,t}$	0.265***			0.416**
	(0.0111)			(0.0639, 0.768)
$\Delta BMI_{n,t}$		0.573***	0.238***	
		(0.168)	(0.0736)	
Year FE	Y	Y	Y	Y
Year-Clustered SE	Y	Y	Y	Y
N	164512	164512	164512	164512
F	568.6	11.68	10.48	13.95
R-Squared	0.0825	0.00747	0.00152	0.0556

Standard errors in parentheses

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

This table displays results for the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \Delta BMI_{n,t} + \boldsymbol{\beta}_1' \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta E_{a,n,t+4h|t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{\beta}_2' \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t,h},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the June change in BMI ($\Delta BMI_{n,t}$). The reduced form regresses quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on $\Delta BMI_{n,t}$. The second stage regresses $\Delta E_{a,n,t+4h|t}$ on instrumented price changes ($\Delta \hat{p}_{a,n,t}$). $\boldsymbol{X}_{n,t}$ includes the log market cap as of the May rank date, the one-year monthly rolling average bid-ask percentage spread, and the banding controls (an indicator for having rank-date market cap in the band including stocks from the Russell 1000 and 2000, an indicator for being in the Russell 2000 in May before reconstitution, and the interaction of these indicators). Column 4 reports the 95% confidence interval for the second-stage coefficient using the tF procedure of Lee et al. (2022). All units are in percentage points (i.e. 1.0 is 1%). The time period is 1999-05:2018-09.

exogenous 1% price increase raises annual EPS expectations by 41 basis points. I report the 95% confidence interval for the second-stage coefficient using the tF procedure of Lee et al. (2022) to address any concerns about how instrument strength impacts second-stage inference. Appendix Figure B1 displays first-stage and reduced-form binscatter plots. Appendix Table B2 displays alternate specifications.

Thus, I reject the null hypothesis of $\alpha = 0$. Prices impact analyst cash flow expectations. These results are inconsistent with models featuring homogeneous analyst and investor cash flow expectations and no feedback from prices to these expectations (as in Section 2.2.1). Yet these results are consistent with models featuring learning from prices due to dispersed information (as in Section 2.2.2) or price extrapolation (as in Section 2.2.3). Hence, it's possible that biases in analyst expectations reflect price variation instead of driving it. We must seriously consider heterogeneous beliefs and endogenous belief formation mechanisms when interpreting these expectations.

Moreover, these revisions of fixed-horizon EPS expectations reflect changes in forecast errors, as discussed in Section 3. Forecast errors increase (in magnitude, i.e. become more negative) in response to exogenous price increases. This result is consistent with the models in Sections 2.2.2 and 2.2.3. This result is not consistent with models in which analysts raise cash flow expectations due to exogenous price increases solely because these increases raise actual future cash flows (e.g. through the relaxation of financial constraints that enables greater investment (Bernanke and Gertler (1986)) or because a firm's managers, customers, or suppliers learn from private information in stock prices and adjust real decisions (Subrahmanyam and Titman (2001); Edmans, Goldstein and Jiang (2015))).

4.2 Economic Magnitude

Is $\alpha = 41$ basis points economically large? As a benchmark, assume analysts use the Gordon growth model:

$$P_t = \frac{\mathbb{E}_t\left[D_{t+1}\right]}{r-g},$$

where $\mathbb{E}_t\left[D_{t+1}\right]$ is expected next-year dividend, and g and r are the growth and discount rates. If analysts adjust only $\mathbb{E}_t\left[D_{t+1}\right]$ so their personal valuations rise to match a 1% price rise, they will raise it by 1%.[26] Thus, $\alpha = 41$ basis points is significant and consistent with analysts raising EPS expectations by less than would fully rationalize price changes (e.g. as in Section 2.2.2 if they view prices as noisy cash-flow signals).

Given the two-stage least squares and OLS estimates (41 and 26 basis points) from Table 2, the proportion of the covariance of prices and analyst expectations this price impact accounts for (2) is 157%. This proportion exceeds 100% because the two-stage least squares estimate exceeds the OLS estimate, which likely reflects "measurement error" in the quarterly price changes $\Delta p_{a,n,t}$. For example, if analysts use price changes only for a subset of days in the quarter (e.g. the one-month price change before the announcement) to update expectations, then the quarterly price change variance exceeds that of the "true" price change analysts respond to $\Delta p^T_{a,n,t}$: $\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right] > \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]$.

Thus, to calculate the true covariance proportion this price impact accounts for, one must multiply (2) by $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right] < 1$. I measure this ratio in Section 5.3 using the method of Pancost and Schaller (2022), which requires one instrument to be applied to multiple outcome variables. I can use this method in Section 5 where I use the flow induced trading (FIT) instrument to measure the impact of prices on analyst expectations of different horizons. The restricted sample in the current section (stocks

[26] Linearizing the Gordon growth model around small percentage changes in price (Δp_t) and next year's expected dividend (Δd_t), and changes in the growth rate (Δg) yields $\Delta p_t \approx \Delta d_t + \left(P_t / \mathbb{E}_t\left[D_{t+1}\right]\right) \Delta g$.

in a narrow window around the reconstitution thresholds in June) renders horizon-specific α estimates noisy (see Appendix Table B3), which leads to underestimation of $\mathbb{V}^{CX}\left[\Delta p^{T}_{a,n,t}\right]/\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ (Pancost and Schaller (2022)). Thus, I calculate the Figure 1 covariance decompositions in Section 5.3 using the FIT instrument. Appendix E reports the decompositions calculated using BMI changes. The covariance decompositions adjusted for measurement error are "conservative" in that they imply the impact of prices on analyst expectations is quantitatively less important than the raw covariance decompositions suggest.

4.3 Robustness

Figure 4 summarizes the robustness checks I conduct for the baseline results in Table 2.

Not winsorizing percentage prices changes, instead of at the baseline 1%, yields similar first-stage (0.637 versus the baseline 0.573) and second-stage ($\alpha = 37.5$ versus the baseline $\alpha = 41.6$ basis points) estimates.

Winsorizing annual EPS expectation revisions from 1% to 9%, instead of the baseline 5%, yields reduced-form estimates from 0.184 to 0.351 (all statistically indistinct from the baseline 0.238), and second-stage α estimates from 32.1 to 61.2 basis points (all statistically indistinct from the baseline 41.6 basis points).

Using alternate bandwidths from 100 to 200 stocks around the cutoffs, instead the baseline 150 stocks, yields similar first-stage (0.534 to 0.593 versus the baseline 0.573), reduced-form (0.211 to 0.229 versus the baseline 0.238), and second-stage ($\alpha = 38.2$ to 40.2 versus the baseline 41.6 basis points) estimates.

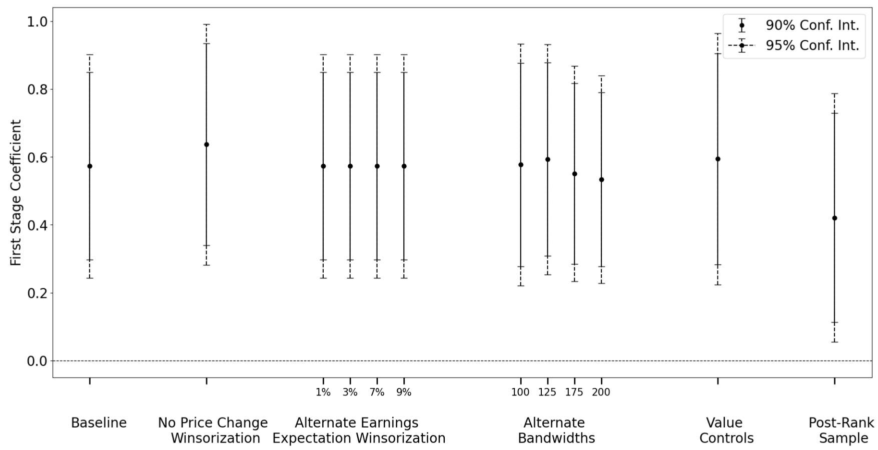
One may be concerned that some of the variation in $\Delta BMI_{n,t}$ may arise from stocks switching from Russell Value to Growth indices. At each May rank date, Russell assigns stocks to Value and Growth indices based on a custom algorithm applied to a proprietary database of analyst forecasts, book-to-price ratio, and sales growth. Thus, it's possible that news for stock n before the rank date could impact analyst beliefs directly, and could impact $\Delta BMI_{n,t}$ by moving stock n from the Value to the Growth indices, or vice versa. This situation threatens exogeneity because I cannot condition on the proprietary information Russell uses to assign stocks to Value and Growth indices.

I address this concern in two ways. First, lacking Russell's proprietary valuation metrics, I control for market-to-book ratio[27] and annual sales growth, and obtain similar first-stage (0.594 versus the baseline 0.573), reduced-form (0.238 versus the baseline 0.228), and second-stage ($\alpha = 38.4$ versus the baseline 41.6 basis points) estimates.

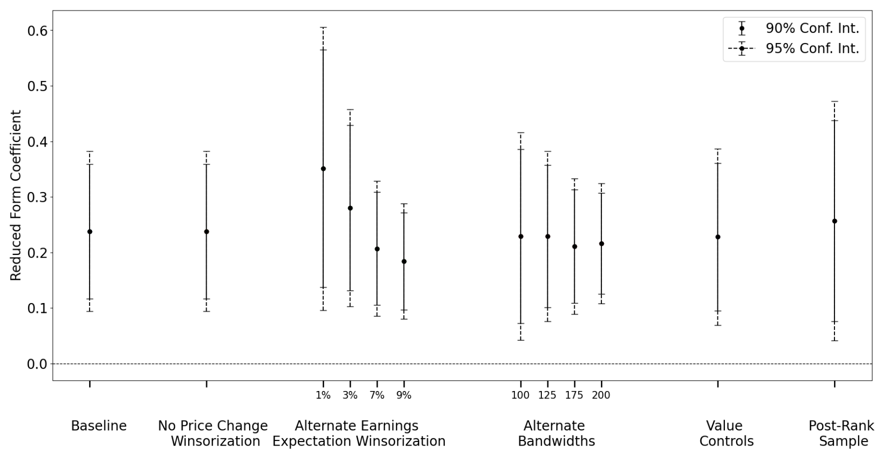
Second, I repeat the analysis using only analyst expectations changes for which the original expectation is

[27] I construct book equity following the approach of Cohen, Polk and Vuolteenaho (2003).

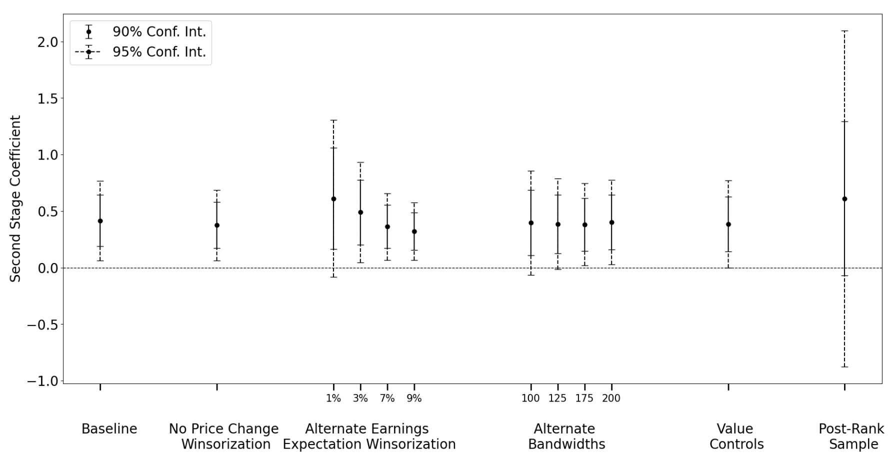
Figure 4: Robustness Checks for Using ΔBMI as Instrument



(a) First-Stage Results



(c) Reduced-Form Results



(c) Second-Stage Results

This figure displays results for alternate specifications of two-stage least squares regression (12). In all specifications $\boldsymbol{X}_{n,t}$ includes the log market cap as of the May rank date, the one-year monthly rolling average bid-ask percentage spread, and the banding controls (an indicator for having rank-date market cap in the band including stocks from the Russell 1000 and 2000, an indicator for being in the Russell 2000 in May before reconstitution, and the interaction of these indicators). In the "Value Controls" specification, $\boldsymbol{X}_{n,t}$ also includes market-to-book ratio and annual sales growth. The solid error bars display 90% confidence intervals, while the dashed error bars display 95% confidence intervals. Panel (c) reports the 90% and 95% confidence intervals for the second-stage coefficients using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1999-05:2018-09.

reported *after the May rank date* but before the end of June, as displayed in Figure 3.[28] Analyst expectations changes in this post-rank sample are not exposed to news from before the rank date that may impact assignment of stocks to the Russell Value and Growth indices, but are still exposed to June reconstitution price changes. This post-rank sample yields similar first-stage (0.421 versus the baseline 0.573) and reduced-form (0.257 versus the baseline 0.228) estimates. The second-stage estimate for this post-rank sample is larger than, but statistically indistinct from, the baseline estimate ($\alpha = 61.1$ versus the baseline 41.6 basis points). Although this post-rank sample second-stage estimate is not statistically significant because the post-rank sample is over ten times smaller than the full sample (11,980 versus 164,512 observations), the first-stage and reduced-form estimates remain significant.

This specification also addresses the concern that BMI increases raise analyst expectations directly instead of through prices. If analysts do not respond to prices but expect passive ownership increases to raise future cash flows due to improved governance or product market outcomes, then these anticipated improvements should appear in the first expectation reported after the May rank date (e.g. May 30) because Russell index assignment depends only on information available at the rank date and can be accurately predicted with public data (see Appendix Table B1). These anticipated improvements should *not* impact expectations *revisions* that occur fully after the rank date (e.g. from June 1 to September 1). Thus, the impact of BMI increases on post-rank date expectations revisions is inconsistent with analysts only attending to passive ownership, but is consistent with analysts responding to prices.

5 Evidence from Mutual Fund Flow-Induced Trading

This section provides evidence that prices impact analyst cash flow expectations using my second exogenous price shock: mutual fund flow-induced trading (FIT).

Stock-level mutual fund trading induced by flows is uninformed: funds tend to scale preexisting holdings proportionally to ex-ante portfolio weights (Frazzini and Lamont (2008)). For example, a $1 inflow induces an S&P 500 index fund to mechanically allocate about five additional cents to Apple, since Apple's weight in the S&P 500 is about 5%. This predicted mechanical component of the cross-sectional trading induced by flows is uninformed.

I use the FIT instrument of Lou (2012) (similar to the flow-to-stock instrument of Wardlaw (2020)).[29]

[28] I obtain rank dates from Ben-David, Franzoni and Moussawi (2019).

[29] The FIT instrument uses predicted trading due to all flows; the flow-to-stock instrument uses only extreme outflows.

I first calculate the quarterly flow to mutual fund i as

$$f_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1} \cdot (1 + \text{Ret}_{i,t})}{\text{TNA}_{i,t-1}}.$$

$TNA_{i,t}$ is fund i's total net assets in quarter t and $\text{Ret}_{i,t}$ is the fund return from quarter $t-1$ to t. The predicted mechanical trading by fund i in stock n induced by this flow is $\text{Shares Held}_{i,n,t-2} \cdot f_{i,t}$. Using the number of shares held from quarter $t-2$ ensures $\text{Shares Held}_{i,n,t-2}$ uses only information available before the change in analyst cash flow expectations from quarter $t-1$ to t. Aggregating across all funds and scaling by shares outstanding yields[30]:

$$\text{FIT}_{n,t} = \sum_{\text{fund } i} \underbrace{\frac{\text{Shares Held}_{i,n,t-2}}{\text{Shares Outstanding}_{n,t-2}}}_{\equiv S_{i,n,t-2}} f_{i,t}. \tag{13}$$

$S_{n,i,t-2}$ is the proportion of all shares of stock n owned by mutual fund i in quarter $t-2$.

Table 3 presents summary statistics. There are $121,553$ analyst-stock-year observations in the matched FIT and LTG expectation sample, spanning 1983 to 2020. There are $3,396,550$ analyst-stock-horizon-year observations in the annual EPS expectation sample, spanning 1982 to 2020 (Appendix Table D4 displays horizon-specific statistics). The availability of I/B/E/S analyst expectations constrains both start points.

I use $\text{FIT}_{n,t}$ as a cross-sectional instrument for price changes. Hence, if analyst belief shocks $\nu_{a,n,t}$ are *cross-sectionally* uncorrelated with $\text{FIT}_{n,t}$ for each analyst a and quarter t

$$\mathbb{E}\left[\text{FIT}_{n,t}\nu_{a,n,t}\right] = 0, \forall a, t, \tag{14}$$

then the FIT instrument satisfies the unconditional exogeneity condition: $\mathbb{E}[\text{FIT}_{n,t}\nu_{a,n,t}] = 0$.

The only source of *cross-sectional* variation in the FIT instrument is the ex-ante ownership shares $S_{i,n,t-2}$. Stocks n for which fund i has greater ownership shares are more exposed to i's flow in this quarter. These stocks have larger magnitudes of flow-induced trading, and so more price pressure. Flows $f_{i,t}$ are at the fund level, and so do *not* create variation across stocks within a quarter. Heterogeneous ownership shares create variation across stocks by creating heterogeneous exposures to flows. Thus, a sufficient condition for

[30] Unlike Lou (2012), but as in Li (2021), I do not multiply the numerator summand by a "partial scaling factor" to reflect the fact that mutual funds may buy or sell less than one dollar in existing positions per dollar of flow due to liquidity or other constraints. Yet while Li (2021) uses the lagged total number of shares held by all mutual funds in the denominator, I use the number of shares outstanding so $\text{FIT}_{n,t} = 0.01$ represents the mutual fund sector buying 1% of stock n's shares.

Table 3: Summary Statistics for FIT Instrument

	$\Delta LTG_{a,n,t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	121553.00	121553.00	121553.00	131333.00
Mean	-0.01	0.03	-0.0001	165.30
Std Dev.	0.04	0.22	0.0039	309.02
Min	-0.12	-0.94	-0.3941	1.00
25%	-0.02	-0.08	-0.0010	44.00
50%	-0.00	0.02	-0.0000	73.00
75%	0.01	0.13	0.0007	132.00
Max	0.10	5.80	0.1845	3712.00

(a) LTG Expectations

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	3396550.00	3396550.00	3396550.00	133902.00
Mean	-0.01	0.03	-0.0001	162.61
Std Dev.	0.19	0.24	0.0042	306.68
Min	-0.59	-0.99	-0.4536	1.00
25%	-0.08	-0.09	-0.0011	43.00
50%	0.00	0.02	-0.0000	72.00
75%	0.07	0.14	0.0008	130.00
Max	0.44	14.05	0.2134	3712.00

(b) Annual EPS Expectations — All Horizons

Summary statistics for quarter-over-quarter changes in LTG expectations $\Delta LTG_{a,n,t}$ (Panel (a)) and revisions in annual EPS expectations for forecast horizons of one to four years $\Delta E_{a,n,t+4h|t}$ (Panel (b)), inter-announcement percentage price changes ($\Delta p_{a,n,t}$), the FIT instrument $\text{FIT}_{n,t}$, and the number of stocks held by mutual funds used to construct the FIT instrument. The first three columns are expressed in absolute terms (i.e. 0.01 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

FIT exogeneity (14) is that the ex-ante ownership shares are exogenous across stocks within each quarter t:

$$\mathbb{E}\left[S_{i,n,t-2}\nu_{a,n,t}\right] = 0, \forall a, i, t. \tag{15}$$

The sufficiency of cross-sectionally exogenous ownership shares is a special case of the result that exogenous shares are sufficient for shift-share instrument exogeneity (Goldsmith-Pinkham, Sorkin and Swift (2020)).

For example, assume there is one fund, analyst, and quarter (so drop subscripts i, a, and t), but N stocks. The FIT instrument is $\text{FIT}_n = S_n f$ ($S_n \neq 1$ because there are other investors). FIT_n is exogenous if and only if the ownership shares are exogenous: $0 = \mathbb{E}\left[\text{FIT}_n \nu_n\right] = \mathbb{E}\left[S_n \nu_n\right] f$, because the flow is constant across stocks. Appendix C Proposition 4 (the same as Goldsmith-Pinkham, Sorkin and Swift (2020) Proposition

2) generalizes this argument.

Exogeneity of the ex-ante ownership shares (15) is plausible because the FIT instrument uses ownership shares from quarter $t-2$, which do not contain information analysts use to update expectations from quarter $t-1$ to t. For example, positive news about Apple in quarter $t-2$ may impact $S_{i,n,t-2}$, but will not be used by analysts to update expectations from quarter $t-1$ to t.

This identification strategy does *not* require flows to be exogenous. Flows may correlate with analyst belief shocks in the time series: $\mathbb{E}\left[f_{i,t}\nu_{a,n,t}\right] \neq 0, \forall a, i, n$. For example, previous work finds correlations of fund flows with surveyed beliefs (Greenwood and Shleifer (2014)), past performance (Ippolito (1992); Chevalier and Ellison (1997); Sirri and Tufano (1998)), past flows (Lou (2012)), and earnings news (Di Maggio et al. (2023)). None of these *time-series* correlations undermines the *cross-sectional* exogeneity of the FIT instrument. (14) can hold even if $\mathbb{E}\left[f_{i,t}\nu_{a,n,t}\right] \neq 0, \forall a, i, n$ because flows do create not cross-sectional variation in $\text{FIT}_{n,t}$ across stocks within a quarter.

For example, one may be concerned that good news about small stocks in quarter t raises analyst expectations for small stocks and drives flows into small-cap funds. This situation *does* threaten exogeneity, but *not* because flows are endogenous. The issue here is the ownership shares are endogenous ((15) fails) because both analyst belief shocks and small-cap fund ownership shares depend on a common stock characteristic: size. Hence, analyst expectations for small stocks are more exposed than those for large stocks to both the price pressure driven by flows into small-cap funds and the good news shock. These correlated exposures to different "aggregate shocks" undermine exogeneity.

Section 5.4 explains the solution is to control for time fixed effects interacted with the problematic stock characteristics (size in this example). Controlling for observed and latent characteristics that explain most ownership share variation yields similar results. Funds holding few stocks raises similar issues, so I construct the FIT instrument only from mutual funds with many holdings and find similar results. Systematic deviations from proportional trading also raise similar issues (Berger (2023)), so I construct the instrument only from passive funds and find similar results.

5.1 Empirical Results

I run the following two-stage least squares regression:

$$
\begin{aligned}
\Delta p_{a,n,t} &= a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t} \\
\Delta y_{a,n,t} &= b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t}.
\end{aligned}
\tag{16}
$$

For analyst institution a and stock n in quarter t, $\Delta y_{a,n,t}$ is either the quarterly LTG expectation change $\Delta\text{LTG}_{a,n,t}$, or the h-year EPS expectation revision $\Delta E_{a,n,t+4h|t}$. $\Delta p_{a,n,t}$ is the corresponding price change between the two quarterly report dates for this analyst institution and stock in quarters $t-1$ and t. FE_t are quarter fixed effects.

Table 4 Panel (a) presents the LTG expectations results.[31] The OLS regression of LTG expectation changes on price changes in column 1 finds a strong association between these objects, as previous work finds (Bordalo et al. (2019, 2022); Nagel and Xu (2021)). The first-stage regression in column 2 is strong with a partial F-statistic (13.3) over the conventional threshold (10): higher flow-induced trading raises prices. The reduced-form coefficient in column 3 is also significant: higher flow-induced trading raises LTG expectations. The second-stage α estimate in column 4 reveals a statistically and economically significant effect of prices on LTG expectations: an exogenous 1% price increase raises LTG expectations by 5 basis points. I report the second-stage 95% confidence interval using the tF procedure of Lee et al. (2022) to address any concerns about how instrument strength impacts second-stage inference.

Panel (b) displays the results for one to four-year horizon annual EPS expectations. The OLS regression of EPS expectation revisions on price changes in column 1 finds a strong association between these objects, as previous work finds (De La O and Myers (2021, 2023)). The first-stage regression in column 2 is strong with a partial F-statistic (16.1) over the conventional threshold (10): higher flow-induced trading raises prices. The reduced-form coefficient in column 3 is also significant: higher flow-induced trading raises annual EPS expectation revisions. The second-stage α estimate in column 4 reveals a statistically and economically significant effect of prices on annual EPS expectation revisions: an exogenous 1% price increase raises annual EPS expectations by 21 basis points.

This $\alpha = 21$ basis points estimate is smaller but statistically indistinct from the $\alpha = 41$ basis points estimate obtained using the ΔBMI instrument in Table 2. These revisions of fixed-horizon EPS expectations reflect changes in forecast errors, as discussed in Section 3, and so this $\alpha > 0$ result is not consistent with models in which analysts raise their expectations due to exogenous price increases solely because these increases raise actual future cash flows.

Thus, I reject the null hypothesis of $\alpha = 0$. There is evidence that prices impact long- and short-term analyst cash flow expectations. These results are inconsistent with models featuring homogeneous analyst and investor cash flow expectations and no feedback from prices (as in Section 2.2.1). Yet they are

[31] Appendix Table D5 displays results with block-bootstrapped (instead of quarterly-clustered) confidence intervals.

consistent with models featuring learning from prices due to dispersed information (as in Section 2.2.2) or price extrapolation (as in Section 2.2.3). Hence, it's possible that biases in analyst expectations reflect price variation instead of driving it. We must seriously consider heterogeneous beliefs and endogenous belief formation mechanisms when interpreting these expectations.

Table 4: Effect of Prices on Analyst Expectations Using FIT as Instrument

	(1)	(2)	(3)	(4)
Panel (a): LTG Expectation Changes				
	OLS	First Stage	Reduced Form	2SLS
$\Delta p_{a,n,t}$	0.0438***			0.0546**
	(0.00209)			(0.0169, 0.0923)
$\text{FIT}_{n,t}$		3.355***	0.183***	
		(0.920)	(0.0602)	
N	121553	121553	121553	121553
F	439.5	13.31	9.242	18.84
R-Squared	0.0399	0.00373	0.000231	
Panel (b): Revisions of Annual EPS Expectations				
$\Delta p_{a,n,t}$	0.279***			0.206**
	(0.0118)			(0.0820, 0.330)
$\text{FIT}_{n,t}$		3.463***	0.714***	
		(0.863)	(0.203)	
N	3396550	3396550	3396550	3396550
F	562.2	16.09	12.32	21.51
R-Squared	0.108	0.00387	0.000229	
Quarter FE	Y	Y	Y	Y
Quarter-Clustered SE	Y	Y	Y	Y

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

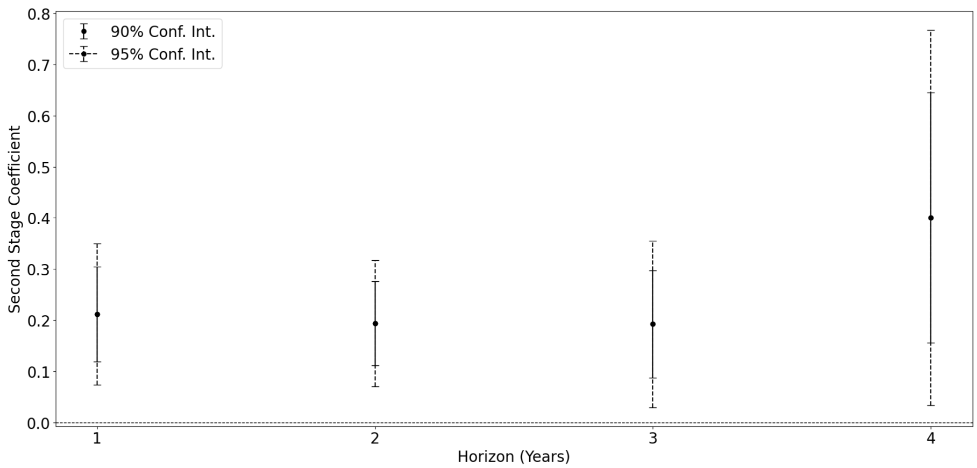
This table displays results for the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \text{LTG}_{a,n,t}$) in Panel (a) and quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) in Panel (b) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). FE_t are quarter fixed effects. All units are in percentage points (i.e. 1.0 is 1%). Column 4 reports the 95% confidence interval for the second-stage coefficients using the tF procedure of Lee et al. (2022). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Appendix Figure D2 displays first-stage and reduced-form binscatter plots. Figures D3 and D4 display alternate specifications. Figure D5 displays results for alternate winsorizations.

Figure 5: Effect of Prices on Term Structure of EPS Expectations



This figure displays the following two-stage least squares regressions results:

$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta E_{a,n,t+4h|t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t},$$

The first stage regresses quarterly percent price changes ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). The second stage regresses changes quarterly revisions to annual EPS expectations ($\Delta E_{a,n,t+4h|t}$) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). I run this regression separately for each horizon h from 1 to 4 years. FE_t are quarter fixed effects. All units are in percentage points (i.e. 1.0 is 1%). The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals, both of which are computed using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. The time period is 1982-04:2020-12.

5.2 Persistence: In the Term Structure and Over Time

The impact of prices on analyst cash flow expectations is persistent in the term structure and time series.

Figure 5 displays the results of two-stage least squares regression (16) for each annual EPS expectation horizon. The α estimates for one to three year horizons range from 19.2 to 21.2 basis points, which are similar to the $\alpha = 20.6$ basis points pooled estimate from Table 4 Panel (b). The four-year estimate of $\alpha = 40.1$ basis points is larger, but is statistically indistinct from the pooled estimate. Appendix Figure D6 displays first-stage and reduced-form results. Thus, this impact appears permanent in the term structure of expectations: it doesn't shrink as forecast horizon grows.

Do analysts revise their expectations ex-post once they have enough public information to determine part of the price change they responded to was driven by noise trading? $\text{FIT}_{n,t}$ uses fund holdings from the end of quarter $t-2$ to instrument the price change spanning parts of quarters $t-1$ and t (as in Figure 2). Funds report these holdings to the SEC with a 60-day delay, so they become public in quarter $t-1$.[32] If

[32]The Thomson Reuters S12 database collects quarterly mutual fund holdings from SEC Forms N-30D, N-Q, and N-CSR, in addition to voluntary mutual fund disclosures (Zhu (2020)). Form N-Q is filed within 60 days of the end of the first and

analysts learn of these filings with a delay (e.g. due to inattention), they would not be able to construct $\text{FIT}_{n,t}$ in real time. If analysts later learn of these filings, construct $\text{FIT}_{n,t}$, realize part of the price change they responded to was noise-driven, and revise their expectations to remove this noise, then the impact of $\text{FIT}_{n,t}$ on analyst expectations will revert over time.

To test for these ex-post revisions, I regress cash flow expectations changes on lagged FIT values:

$$\Delta y_{a,n,t} = \sum_{s=0}^{L} \beta_s \text{FIT}_{n,t-s} + FE_t + FE_n + e_{a,n,t}.$$

$\Delta y_{a,n,t}$ is the LTG expectation change or annual EPS expectation revision. $L = 0, \ldots, 4$ is the max lag.

Figure 6 displays the sum of the coefficients on the lagged FIT instruments for each maximum lag $L = 0, \ldots, 4$: $\sum_{s=0}^{L} \beta_s$. The vertical line at the one quarter lag highlights that the first "true lag" is actually quarter $t-2$, since $\text{FIT}_{n,t-1}$ overlaps with $\Delta y_{a,n,t}$ (as shown in Figure 2). At horizons of up to one year, I find no evidence that analysts revise their LTG or one to four year EPS expectations to remove their responses to noise-drive price changes. Thus, the impact of prices on analyst cash flow expectations is persistent.

Appendix Figure D7 presents some evidence that this impact reverts at longer horizons (twelve to sixteen quarters) once the FIT instrument's price impact starts to revert (Lou (2012) finds it reverts at similar horizons of eight to sixteen quarters). If prices impact analyst expectations, then analysts should lower their expectations due to these ex-post price reversions. I find some evidence of this prediction, but lack power at longer lags for a sharp statistical test.

5.3 Economic Magnitude

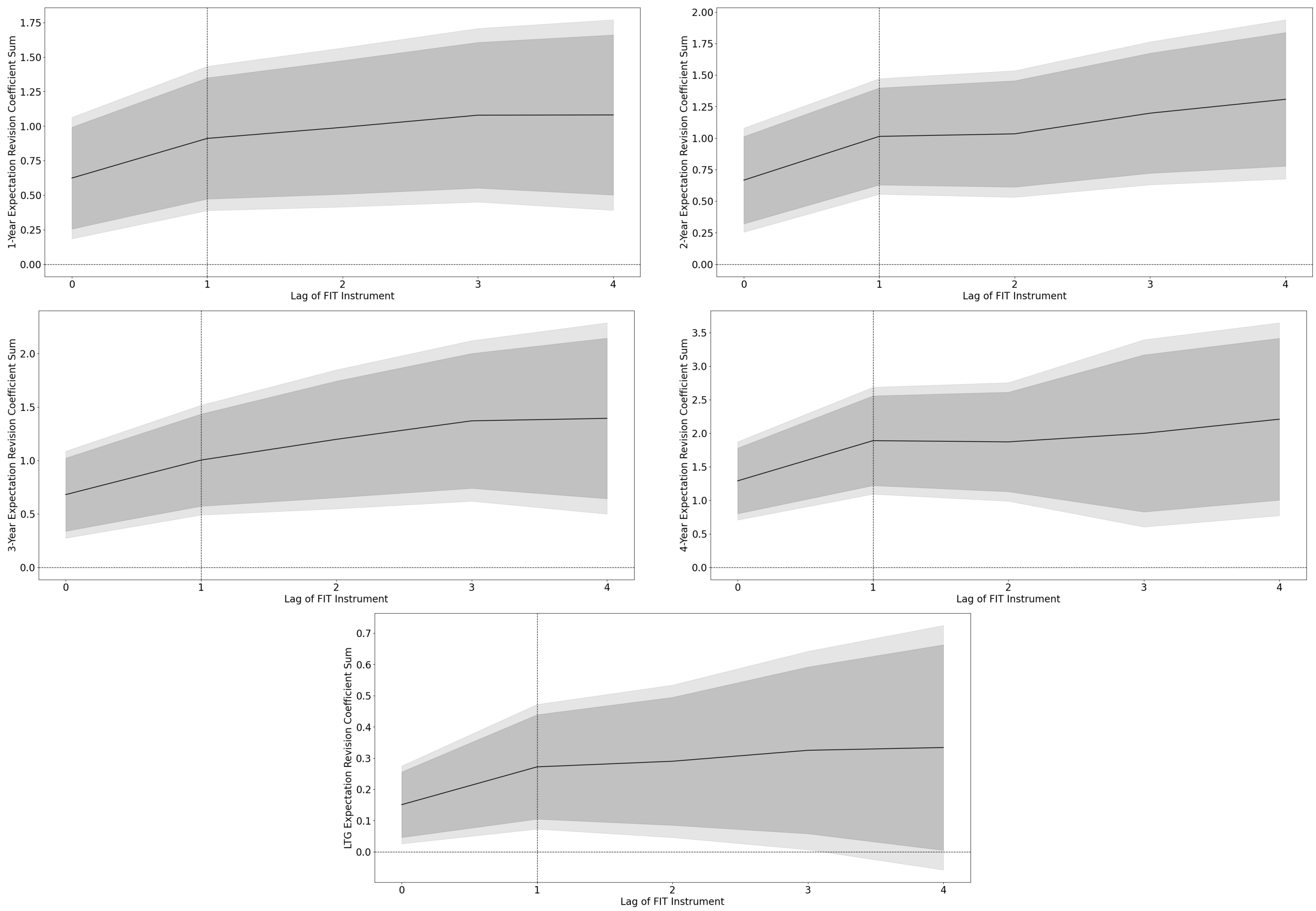
Are the price impacts in Table 4 economically large? Under the Gordon growth model, if analysts adjust only long-run growth rates so their personal valuations rise to match a 1% price rise, they will raise g by $\mathbb{E}_t [D_{t+1}] / P_t\%$, which is about 4 basis points on average for the aggregate market.[33] Thus, if LTG expectations proxy for analyst beliefs about g, $\alpha = 5$ basis points is consistent with analysts revising LTG expectations to fully rationalize price changes.

As in Section 4.2, if analysts adjust only $\mathbb{E}_t [D_{t+1}]$ to match a 1% price rise, they will raise it by 1%.

third quarters of the fund's fiscal year (Securities and Commission (2023*b*)). Form N-CSR is filed within 10 days of the fund's "transmission to stockholders of any annual or semi-annual report that is required to be transmitted to stockholders pursuant to Rule 30e-1 under the [Investment Company Act of 1940]" (Securities and Commission (2023*a*)). According to the Act, these annual and semi-annual reports are made within 60 days of the fund's fiscal half-year end and fiscal year-end, i.e. the second and fourth quarters of the fund's fiscal year (Office (2023)).

[33] Linearizing the Gordon growth model around small percentage changes in price (Δp_t) and next year's expected dividend (Δd_t), and changes in the growth rate (Δg) yields $\Delta p_t \approx \Delta d_t + (P_t / \mathbb{E}_t [D_{t+1}]) \Delta g$.

Figure 6: Lack of Immediate Reversal in Impact of FIT on Analyst Cash Flow Expectations



This figure displays the coefficient sums $\sum_{s=0}^{L} \beta_s, L = 1, \ldots, 4$, from the following regression:

$$\Delta y_{a,n,t} = \sum_{s=0}^{L} \beta_s \text{FIT}_{n,t-s} + FE_t + FE_n + e_{a,n,t}.$$

$\Delta y_{a,n,t}$ is either the quarter-over-quarter change in LTG expectations for analyst institution a for stock n in quarter t $\Delta \text{LTG}_{a,n,t}$, or the quarterly revision to annual EPS expectations with horizons of one to four years $\Delta E_{a,n,t+4h|t}$. FE_t and FE_n are quarter and stock fixed effects. Dark and light shaded areas represent 90% and 95% confidence intervals, respectively. Standard errors are clustered by quarter and stock. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Thus, $\alpha = 21$ basis points is significant and consistent with analysts raising EPS expectations by less than would fully rationalize price changes (e.g. as in Section 2.2.2 if they view prices as noisy cash-flow signals).

Given the two-stage least squares and OLS estimates (5 or 21 and 4 or 28 basis points) from Table 4, the proportion of the covariance of prices and analyst expectations this price impact accounts for (2) is 125% and 75% for the LTG and annual EPS expectations. As in Section 4.2, these are likely overestimates due to "measurement error": quarterly price changes $\Delta p_{a,n,t}$ may not be the "true" price changes $\Delta p^T_{a,n,t}$ analysts respond to (e.g. analysts may respond only to the price change one month before the announcement). Thus, the true covariance proportion is

$$\frac{\alpha \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]}{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y_{a,n,t}\right)} = \frac{\text{Two Stage Least Squares Estimate of } \alpha}{\text{OLS Coefficient in Regression of } \Delta y_{a,n,t} \text{ on } \Delta p_{a,n,t}} \cdot \underbrace{\frac{\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}}_{\leq 1}. \tag{17}$$

I use the method of Pancost and Schaller (2022) to measure $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$. Let α_h and α_h^{OLS} be the two-stage least squares and OLS estimates for horizon h (LTG or one to four years). If α_h^{OLS} has omitted variable bias (due to common information or sentiment shocks impacting both analyst expectations directly and prices via investor expectations) and measurement error, it is a linear function of α_h:

$$\alpha_h^{OLS} = \frac{\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}\alpha_h + OVB_h,$$

Pancost and Schaller (2022) demonstrate that in an OLS regression of α_h^{OLS} on α_h

$$\alpha_h^{OLS} = a + \theta\alpha_h + e_h, \tag{18}$$

θ consistently estimates $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$.[34] $\alpha_h / \alpha_h^{OLS} \cdot \theta$ recovers corrected proportion (17).[35]

Figure 1 displays the estimated corrected covariance proportion (17) for the LTG and pooled set of annual EPS expectations in quarterly block-bootstrapped samples. The impact of prices on analyst cash flow expectations explains 61.0% and 39.7% of the cross-sectional covariance of prices with LTG and annual EPS expectations. Common information or sentiment shocks to analysts and investors explain the remainder.

[34] In this regression, I constrain θ so that $\theta \leq 1$ and the corrected covariance proportion (17) does not exceed 1 for any horizon (which is a tighter constraint). θ can be arbitrarily small. The point estimates for the corrected covariance proportion are similar using unconstrained regressions, but the upper end of the confidence interval is wider (see Appendix Table E8). Estimation error in α_h attenuates θ, creating underestimation of $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ (Pancost and Schaller (2022)).

[35] If $\Delta p_{a,n,t}$ has non-classical measurement error, then θ does not consistently estimate $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$, but $\alpha_h / \alpha_h^{OLS} \cdot \theta$ still recovers corrected covariance proportion (17), as discussed in Appendix E.1.

Appendix E provides estimation details, estimates of θ, and covariance decompositions for all horizons.

5.4 Threat to Exogeneity: Common Characteristics

The main threat to FIT exogeneity is that both ex-ante ownership shares and analyst belief shocks depend on common stock characteristics. In this case, (15) fails: the ownership shares are not exogenous. As discussed above, FIT exogeneity is *not* threatened by flows chasing returns and stock characteristics, or correlating with surveyed beliefs and past flows because ownership share exogeneity is sufficient for FIT exogeneity.

If ownership shares $S_{i,n,t-2}$ and analyst belief shocks $\nu_{a,n,t}$ depend on common stock characteristics, they correlate across stocks within a quarter. For example, small-cap funds have larger ownership shares in small stocks than in large stocks. At the same time, a small-firm tax cut this quarter raises analysts expectations more for small stocks than large stocks. Thus, stocks with higher small-cap fund ownership shares have higher analyst belief shocks, which violates (15).

To be concrete, consider this factor structure in ownership shares and analyst belief shocks[36]:

$$S_{i,n,t-2} = \boldsymbol{c}_i' \boldsymbol{X}_n + \tilde{S}_{i,n,t-2} \tag{19}$$

$$\nu_{a,n,t} = \boldsymbol{X}_n' \boldsymbol{\eta}_t + \tilde{\nu}_{a,n,t}. \tag{20}$$

In (19), fund i's ownership shares $S_{i,n,t-2}$ depend cross-sectionally on stock characteristics $\boldsymbol{X}_n$ (small-cap funds have larger ownership shares in small firms). In (20), the impact of aggregate shocks $\boldsymbol{\eta}_t$ (the tax news) on analyst beliefs depends on characteristics (size). $\tilde{S}_{i,n,t-2}$ and $\tilde{\nu}_{a,n,t}$ are uncorrelated with other objects.

In this case, the ownership shares are *not* cross-sectionally exogenous ((15) fails):

$$\forall a, i, t : \mathbb{E}\left[S_{i,n,t-2}\nu_{a,n,t}\right] = \boldsymbol{c}_i' \mathbb{E}\left[\boldsymbol{X}_n \boldsymbol{X}_n'\right] \boldsymbol{\eta}_t \neq 0. \tag{21}$$

Since the ownership shares are not cross-sectionally exogenous, neither is $\text{FIT}_{n,t}$ ((14) fails):

$$\text{FIT}_{n,t} = \sum_i f_{i,t} S_{i,n,t-2} = \underbrace{\left(\sum_i \boldsymbol{c}_i f_{i,t}\right)'}_{\equiv \boldsymbol{\beta}_t'} \boldsymbol{X}_n + \underbrace{\sum_i \tilde{S}_{i,n,t-2} f_{i,t}}_{\equiv \tilde{\text{FIT}}_{n,t}} \tag{22}$$

$$\forall a, t : \mathbb{E}\left[\text{FIT}_{n,t}\nu_{a,n,t}\right] = \boldsymbol{\beta}_t' \mathbb{E}\left[\boldsymbol{X}_n \boldsymbol{X}_n'\right] \boldsymbol{\eta}_t \neq 0.$$

[36] A more general specification $\nu_{a,n,t} = \boldsymbol{\lambda}_{a,n}' \boldsymbol{\eta}_t + \tilde{\nu}_{a,n,t}$, $\boldsymbol{\lambda}_{a,n} = \boldsymbol{\Gamma}_a \boldsymbol{X}_n + \tilde{\boldsymbol{\lambda}}_{a,n}$ does not impact any of the arguments.

Simply put, cross-sectional FIT variation comes from heterogeneous flow exposures (i.e. heterogeneous ownership shares) that correlate with other shock exposures: small stocks are more exposed to small-cap fund flows and the tax news.

The solution is to control for these common stock characteristics interacted with quarter indicators. Doing so removes the FIT variation driven by these characteristics ($\boldsymbol{\beta}_t' \boldsymbol{X}_n$). From (22), using $\text{FIT}_{n,t}$ while linearly controlling for characteristics interacted with quarter indicators is the same as constructing the FIT instrument from residual ownership shares $\tilde{S}_{i,n,t-2}$: $\text{F}\tilde{\text{I}}\text{T}_{n,t} = \sum_i \tilde{S}_{i,n,t-2} f_{i,t}$. $\text{F}\tilde{\text{I}}\text{T}_{n,t}$ is exogenous since the residual ownership shares are.[37]

Which characteristics should one control for? While I cannot identify which characteristics correlate with analyst belief shocks (in general one cannot test an exogeneity condition), I can identify those that explain significant cross-sectional variation in ownership shares. If controlling for the most important determinants of ownership shares has little impact on estimates of α, that suggests this common characteristics concern does not prove serious empirically.

Section 5.4.1 controls for stock characteristics associated with the investment styles of funds that drive most FIT instrument variation. Since ownership shares and analyst belief shocks may also depend on unobserved characteristics, Section 5.4.2 controls for latent characteristics from a latent factor model that explain most cross-sectional ownership share variation. Funds holding few stocks raises similar issues: ownership shares depend on a specific stock characteristic: firm identity. Section 5.4.3 constructs the FIT instrument from only funds with many holdings. Systematic deviations from proportional trading raise similar issues (Berger (2023)). Section 5.4.4 constructs the FIT instrument from only passive funds, which generally adhere to proportional trading.

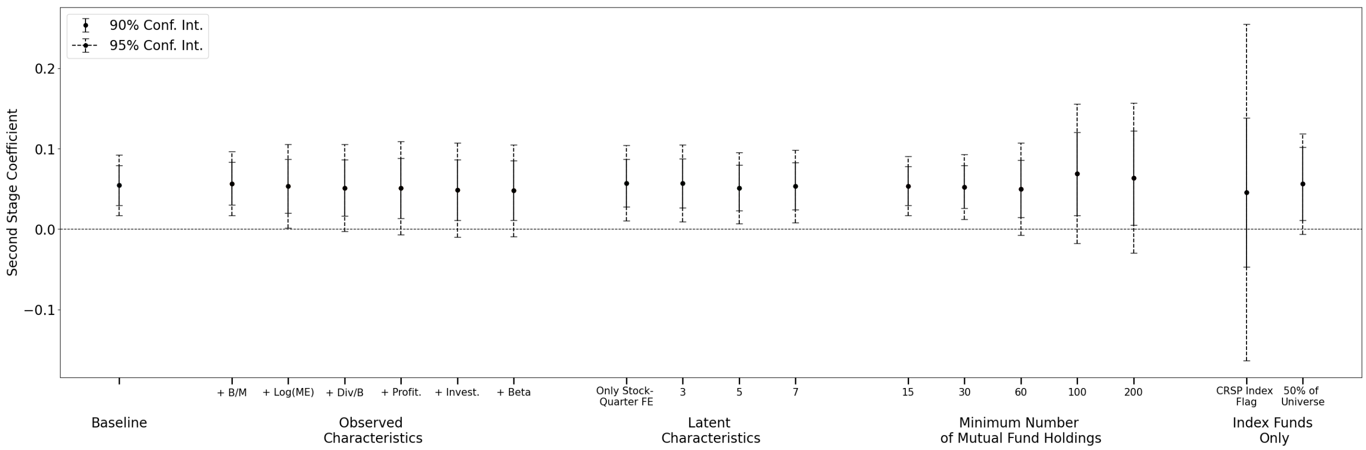
As summarized in Figure 7, these alternate specifications yield α estimates similar to the baseline estimates in Table 4. Appendix Figures D8 and D9 present the first-stage and reduced-form results.

5.4.1 Controlling for Observed Stock Characteristics

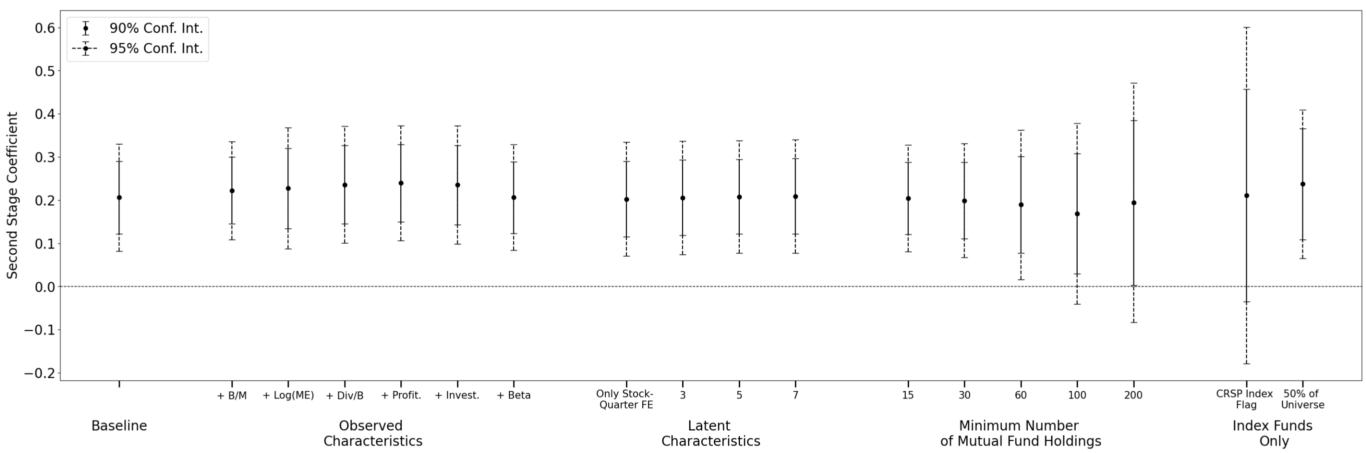
One of the most important determinants of within-fund cross-sectional ownership share variation is investment style. Small-cap funds have larger ownership shares in small stocks than in large stocks. Value funds have larger ownership shares in value stocks than in growth stocks. Thus, characteristics associated with fund styles explain much ownership share variation, and so may threaten share exogeneity. I use CRSP

[37] $\text{F}\tilde{\text{I}}\text{T}_{n,t}$ is exogenous ($\mathbb{E}\left[\text{F}\tilde{\text{I}}\text{T}_{n,t}\nu_{a,n,t}\right] = 0, \forall a, t$) since $\tilde{S}_{i,n,t-2}$ are exogenous ($\mathbb{E}\left[\tilde{S}_{i,n,t-2}\nu_{a,n,t}\right] = 0, \forall a, i, t$).

Figure 7: Robustness Checks for Using FIT Instrument



(a) LTG Expectations



(b) Annual EPS Expectations

This figure displays the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{\beta}_2^{'} \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t}.$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$ in Panel (a)) and quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$ in Panel (b)) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). $\boldsymbol{X}_{n,t}$ includes quarter indicators interacted with either observed (in the "Observed Characteristics" specifications) or latent (in the "Latent Characteristics" specifications) stock characteristics. The observed stock characteristics include: book-to-market ratio, log market equity, dividend-to-book equity ratio, profitability, investment, and market beta. For the "Observed Characteristics" specifications , each subsequent column adds an additional control variable (e.g. the right-most column represents the results of the regression with all six control variables). The time period for the "Observed Characteristics" and "Latent Characteristics" specifications is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions. For the "Minimum Number of Mutual Fund Holdings" specifications, each column constructs $\mathrm{FIT}_{n,t}$ only from mutual funds that have at least M holdings, where M is labeled on the x-axis. The time period for the "Minimum Number of Mutual Fund Holdings" specifications is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions. The "Index Funds Only" specifications construct $\mathrm{FIT}_{n,t}$ from only mutual funds identified as index funds by one of the two criteria listed on the x-axis. The time period for the "Index Funds Only" specifications is 1984-09:2020-12. FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals, both of which are computed using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%).

Table 5: Rotemberg Weights as Percentage of Total for Most Important Fund Styles

Style	LTG	Style	Annual EPS Expectation
Domestic Equity Growth	82.74	Domestic Equity Small Cap	38.08
Domestic Equity Small Cap	27.23	Domestic Equity Growth	35.97
Domestic Equity Income	3.98	Domestic Equity Mid Cap	10.20
Domestic Equity Growth & Income	2.37	Domestic Equity Micro Cap	4.08

Total Rotemberg weights (expressed as percentages of the total weight) for top five most important fund styles for both the LTG expectations ($\Delta LTG_{a,n,t}$) and annual EPS expectations ($\Delta E_{a,n,t+4h|t}$) samples.

style codes to identify fund styles.[38]

Which fund styles are most important to control for? As a shift-share instrument, the FIT instrument is equivalent to using ownership shares (interacted with quarter indicators) as cross-sectional instruments for price changes in an over-identified system with a specific GMM weighting matrix (Goldsmith-Pinkham, Sorkin and Swift (2020)). Implicitly, some funds, and so some styles, receive more weight in this estimation. I calculate these "Rotemberg weights" from Goldsmith-Pinkham, Sorkin and Swift (2020) for each fund-quarter, and then aggregate to the style level (see Appendix D.1 for details). Violation of share exogeneity (15) for higher Rotemberg weight styles biases the α estimates more. Thus, characteristics associated with high-weight styles are the most important to control for.

As displayed in Table 5, the most important styles for both the LTG and annual EPS expectations samples are cap-based and growth/income-based. The top five styles account for 116% and 88% of the total Rotemberg weight for these samples (a few styles have slightly negative weights, as detailed in Appendix D.1). These styles drive most of the FIT instrument variation. For example, since the most important style for the LTG expectations sample is growth, much of the the FIT instrument variation comes from comparing stocks that growth funds have large ownership shares in (growth stocks) to stocks these funds have small ownership shares in (value stocks). Analyst belief shocks for growth and value stocks having heterogeneous exposures to other aggregate shocks may violate share exogeneity. Since cap-based fund ownership shares correlate with size and growth/income-based fund ownership shares correlate with valuation metrics, these stock characteristics may threaten share exogeneity and should be controlled for.

I control for log market equity as a size measure, book-to-market and dividend-to-book ratios as valuation metrics, and profitability, investment, and market beta as popular characteristics that may correlate with ownership shares.[39] These six characteristics explain 46% of within fund-quarter ownership share

[38]CRSP style codes are defined in this document: https://wrds-www.wharton.upenn.edu/documents/1303/MFDB_Guide.pdf.

[39]When using change in analyst expectations $\Delta y_{a,n,t}$, I use characteristics from quarter $t-2$, which is the same quarter the ownership shares $S_{i,n,t-2}$ are taken from to construct $\text{FIT}_{n,t}$. Profitability is the ratio of operating profits over book equity.

variation (see Appendix D.2 for details).

As Figure 7 displays, controlling for these characteristics interacted with quarter indicators does not affect the α estimates. Sequentially adding controls, the LTG expectations changes estimates in Panel (a) range from $\alpha = 4.8$ to 5.7 basis points, which are close to the baseline $\alpha = 5.5$ basis points estimate from Table 4 Panel (a). The annual EPS expectations revisions estimates in Panel (b) range from $\alpha = 20.6$ to 24.0 basis points, which are close to the baseline $\alpha = 20.6$ basis points estimate from Table 4 Panel (b).

5.4.2 Controlling for Latent Stock Characteristics

While controlling for characteristics associated with the most important fund styles does not impact the α estimates, ownership shares and analyst belief shocks may both depend on unobserved characteristics. To address this concern, I fit a latent factor model to estimate the latent characteristics that explain cross-sectional ownership share variation:

$$S_{i,n,t} = \boldsymbol{c}_{i,t}^{'}\boldsymbol{X}_{n,t} + FE_{i,t} + FE_{n,t} + \tilde{S}_{i,n,t}. \tag{23}$$

I fit latent factor model (23) to the fund $\times$ stock panel in each quarter using regularized singular value decomposition (Funk (2006)) (details in Appendix D.3).[40] The stock-quarter fixed effect and the first seven characteristics explain 75% of within fund-quarter ownership share variation (details in Appendix D.2).

As Figure 7 displays, controlling for these eight latent characteristics interacted with quarter indicators does not affect the α estimates. The LTG expectations changes estimates in Panel (a) range from $\alpha = 5.1$ to 5.7 basis points, which are similar to the baseline $\alpha = 5.5$ basis points estimate in Table 4 Panel (a). The annual EPS expectations revisions estimates in Panel (b) range from $\alpha = 20.3$ to 20.9 basis points, which are similar to the baseline $\alpha = 20.6$ basis points estimate in Table 4 Panel (b). Appendix Figures D10, D11, and D12 display the first-stage, reduced-form, and second-stage results for all numbers of latent factors.

5.4.3 Requiring Minimum Number of Holdings

Funds holding few stocks may violate share exogeneity. In the extreme case where each fund holds one of N total stocks, any stock-specific shocks (e.g. earnings surprises) will violate share exogeneity. For example,

Investment is log annual growth rate of assets. Market beta is constructed from 60-month rolling regressions using returns in excess of one-month Treasury bill rates. I winsorize profitability, investment, and market beta at the 2.5th and 97.5th percentiles. Since dividends and book equity are non-negative, I winsorize them at the 97.5th percentile.

[40] Given the sparsity of the data (most funds do not hold most stocks), I use L2 (i.e. ridge) regularization to estimate the factor model more efficiently. Regularization biases the factor and loading estimates toward zero to reduce their variance.

only analyst expectations for Apple are exposed to flows into the "Apple fund" and Apple's earnings surprise.

In this example, ownership shares and analyst belief shocks depend on a specific characteristic: firm identify. $\boldsymbol{X}_n$ in (19) and (20) is an N-dimensional vector of stock indicators.: $\boldsymbol{X}_n = [1_{j=n}]_{j=1}^{N}$. Since both ownership shares and analyst belief shocks depend on a common characteristic, share exogeneity (15) is violated, as in (21). The previous characteristic controls do not remove variation due to firm identity: one would have to control for stock-quarter fixed effects, which would absorb all FIT instrument variation.

This fund concentration concern is not serious empirically. In my sample the average (median) number of stocks held by each fund in each quarter is over 160 (70) (see Table 3). Moreover, strong fund concentration would create a weak factor structure in ownership shares: spanning them would require N (the number of stocks) characteristics. Yet I find a strong factor structure: six observed characteristics explain 46% of ownership share variation (in Section 5.4.1); eight latent characteristics explain 75% (in Section 5.4.2).

To further address this concern, I construct alternate $\text{FIT}_{n,t}$ versions using only funds with at least $M \in [15, 200]$ holdings in quarter t. As Figure 7 displays, doing so does not affect the α estimates. The LTG expectations changes estimates in Panel (a) range from $\alpha = 5.0$ to 6.9 basis points, which are similar to the baseline $\alpha = 5.5$ basis points estimate in Table 4 Panel (a). The annual EPS expectations revisions estimates in Panel (b) range from $\alpha = 16.8$ to 20.4 basis points, which are similar to the baseline $\alpha = 20.6$ basis points estimate in Table 4 Panel (b). Power decreases as the minimum number of holdings rises because excluding some funds reduces the FIT instrument's variation.

5.4.4 Using only Passive Funds

Berger (2023) finds systematic deviations from the proportional trading assumption (that funds scale holdings proportionally in response to flows) can create selection bias. Per Berger (2023), while the proportional trading assumption may normally hold (as shown by Lou (2012)), it may not hold when funds face extreme outflows (e.g. due to liquidity costs). In the context of the Edmans, Goldstein and Jiang (2012) $MFFLOW$ instrument, Berger (2023) argues these systematic deviations occur for stocks with certain characteristics because this instrument uses only the predicted trading driven by extreme outflows.[41] Since the Lou (2012) FIT instrument I use includes the predicted trading driven by *all* flows (not just extreme outflows), these systematic deviations are less likely in my setting.

Still, I take this concern seriously, because systematic deviations from proportional trading would create

[41] Berger (2023) raises similar concerns for the flow-to-stock instrument from Wardlaw (2020).

a dependence of ownership shares on stock characteristics, which threatens share exogeneity (15).[42] To address this concern, I exploit the observation from Berger (2023) that the proportional trading assumption generally holds for passive funds, which maintain portfolios close to their benchmarks. Even if passive funds do not perfectly adhere to the proportional trading assumption, they deviate much less from it than active funds.[43] Thus, if selection bias drives the baseline results in Table 4, then constructing the FIT instrument from only passive funds should yield much weaker results. Yet doing so yields similar results, which suggests this selection bias concern is not serious empirically.

I use two definitions of "passive funds". First, following Berger (2023) I classify all funds with the CRSP index fund flag[44] or "target date" in their names as passive. As Figure 7 displays, constructing the FIT instrument from only these passive funds yields α estimates of 4.6 and 20.1 basis points for the LTG and annual EPS expectations samples, which are similar to the baseline 5.5 and 20.6 basis points estimates from Table 4. While these point estimates are similar, they are not statistically significant because the CRSP index fund flag is only available since 1998 and the first named "target date" funds appear in my sample in 2005, so the sample sizes are much smaller.[45]

To increase power, I classify funds as passive based on the proportion of their investment universes they hold. Similar to Koijen and Yogo (2019), I define the investment universe of a fund as the set of all stocks it has ever held in the last five years. A fund is passive if it holds at least $M\%$ of the stocks in its universe.[46] For $M = 50\%$, this rule successfully classifies 84% of fund-quarter observations as passive or not based on the CRSP index fund flag and "target date" name definition, with a false positive rate of only 8%. As Figure 7 displays, constructing the FIT instrument from only funds holding least 50% of their universes yields α estimates of 5.6 and 23.7 basis points for the LTG and annual EPS expectations samples, which are similar to the baseline 5.5 and 20.6 basis points estimates. Using thresholds above 50% improves classification

[42] Formally, let predicted mechanical trading by fund i in stock n due to flow $f_{i,t}$ be Shares $\text{Held}_{i,n,t-2} \cdot f_{i,t}$. Aggregating across mutual funds yields $\text{FIT}_{n,t}$ from (13). Let the actual trading be $\psi_{i,n,t} \cdot f_{i,t}$, and define $\text{FIT}^{\psi}_{n,t} = \sum_{\text{fund } i} S^{\psi}_{i,n,t} f_{i,t}$, where $S^{\psi}_{i,n,t} = \psi_{i,n,t} / \text{Shares Outstanding}_{n,t}$. If actual trading systematically deviates from proportional trading depending on stock characteristics $\boldsymbol{X}_{n,t}$ (Shares $\text{Held}_{i,n,t-2} = \psi_{i,n,t} + \boldsymbol{c}' \boldsymbol{X}_{n,t}$), then deviations of the ownership shares the FIT instrument uses ($S_{i,n,t}$) from the shares that "should" be used ($S^{\psi}_{i,n,t}$) depend on stock characteristics: $S_{i,n,t} = \left(S^{\psi}_{i,n,t} \cdot \text{Shares Outstanding}_{n,t} + \boldsymbol{c}' \boldsymbol{X}_{n,t} \right) / \text{Shares Outstanding}_{n,t-2}$.

[43] Heath et al. (2022) note that passive ETFs do not hold all of the stocks in their benchmarks, but rather hold a sample that optimally trades off tracking error versus transactions costs. Still, Heath et al. (2022) find that passive ETFs hold the vast majority of stocks in their benchmarks (97%, 89%, and 63% of stocks in the top, middle, and bottom liquidity terciles).

[44] Specifically, I include only funds with the "D" flag, which stands for "pure index fund".

[45] For the LTG and annual EPS expectations samples, using the FIT instrument constructed from funds with the CRSP index fund flag and target date funds yields 80,870 (down from 121,553) and 2,392,080 (down from 3,396,550) observations.

[46] Specifically, I apply this classification at the fund-quarter level.

accuracy and yields similar results, at the expense of power, as Appendix Figure D15 displays.[47] Appendix Figures D13 and D14 display the first-stage and reduced-form results.

6 Conclusion

Prices impact analyst cash flow expectations. Using instruments based on Russell index reconstitution and mutual fund flow-induced trading, I provide the first evidence that analysts raise their long- and short-term cash flow expectations in response to exogenous price increases unrelated to fundamentals. This price impact explains approximately half the covariance between prices and analyst cash flow expectations.

This price impact has important implications for how asset pricing and macro-finance researchers interpret analyst expectations. My results are inconsistent with typical models used to interpret analyst cash flow expectations, which feature homogeneous analyst and investor cash flow expectations, and no feedback from prices to these expectations. Yet these results are consistent with models that relax either of these two assumptions. On one hand, there are models that allow investor and analyst cash flow expectations to differ, such as models with dispersed information in which analysts learn from prices that reflect investors' private information. In these models, at least some investors' cash flow expectations do not depend on prices, but then analysts have different expectations from these investors. On the other hand, there are models that allow common investor and analyst cash flow expectations to depend on prices, such as models of price extrapolation. In these models, exogenous price increases today raise expectations of future prices, and agents mechanically raise *cash flow* expectations to justify these higher *price* expectations.

Compared to typical models, these latter two model classes present different interpretations of correlations of analyst expectations with prices, future returns, and forecast errors. In models in previous work, biases in analyst cash flow expectations are shared by investors and so distort prices. In these alternative models, analyst expectations may only *reflect* price distortions from other variables (e.g. noise shocks, discount rates, biased price expectations). My results imply these alternative interpretations should be taken seriously when working with analyst cash flow expectations, and so motivate further study of heterogeneous beliefs and endogenous belief formation mechanisms.

Moreover, these results raise important questions about how investors and analysts form beliefs and the extent to which analyst beliefs proxy for investor beliefs. Are analyst cash flow expectations a good proxy

[47] Classification accuracy based on the CRSP index fund flag and "target date" name definition of "passive" for the 50% of Universe definition is 84% with a false positive rate of 8%. Accuracy for the 60% of Universe definition is 87% with a false positive rate of 3%. Accuracy for the 70% of Universe definition is 88% with a false positive rate of 1%.

for the beliefs of a large, price-relevant group of investors? What belief formation mechanism creates this price impact on cash flow expectations? Do large groups of investors share this mechanism with analysts? Do investors have biased cash flows expectations? Are these biased expectations important drivers of asset prices? Given the impact of prices on analyst cash flow expectations, these expectations alone are likely insufficient to answer these questions. Instead, direct measures of investor beliefs or empirical strategies that account for belief heterogeneity will likely prove necessary.

References

Albert Jr, Robert L, and Timothy R Smaby. 1996. "Market response to analyst recommendations in the "dartboard" column: the information and price-pressure effects." *Review of Financial Economics*, 5(1): 59–74.

Appel, Ian R, Todd A Gormley, and Donald B Keim. 2016. "Passive investors, not passive owners." *Journal of Financial Economics*, 121(1): 111–141.

Appel, Ian R, Todd A Gormley, and Donald B Keim. 2019. "Standing on the shoulders of giants: The effect of passive investors on activism." *The Review of Financial Studies*, 32(7): 2720–2774.

Appel, Ian, Todd A Gormley, and Donald B Keim. 2021. "Identification using Russell 1000/2000 index assignments: A discussion of methodologies." *Critical Finance Review, Forthcoming.*

Asquith, Paul, Michael B Mikhail, and Andrea S Au. 2005. "Information content of equity analyst reports." *Journal of financial economics*, 75(2): 245–282.

Bai, Jushan, and Serena Ng. 2019. "Rank regularized estimation of approximate factor models." *Journal of Econometrics*, 212(1): 78–96. Big Data in Dynamic Predictive Econometric Modeling.

Barber, Brad M, and Douglas Loeffler. 1993. "The "Dartboard" column: Second-hand information and price pressure." *Journal of Financial and Quantitative Analysis*, 273–284.

Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman. 2001. "Can investors profit from the prophets? Security analyst recommendations and stock returns." *The Journal of Finance*, 56(2): 531–563.

Bastianello, Francesca, and Paul Fontanier. 2021*a*. "Partial Equilibrium Thinking, Extrapolation, and Bubbles."

Bastianello, Francesca, and Paul Fontanier. 2021*b*. "Partial equilibrium thinking in general equilibrium." Working paper, Harvard University.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. 2019. "An improved method to predict assignment of stocks into Russell indexes." National Bureau of Economic Research.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song. 2022. "Ratings-driven demand and systematic price fluctuations." *The Review of Financial Studies*, 35(6): 2790–2838.

Berger, Elizabeth A. 2023. "Selection bias in mutual fund fire sales." *Journal of Financial and Quantitative Analysis*, 58(3): 1039–1077.

Bernanke, Ben S, and Mark Gertler. 1986. "Agency costs, collateral, and business fluctuations."

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019. "Diagnostic expectations and stock returns." *The Journal of Finance*, 74(6): 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, Rafael LaPorta, and Andrei Shleifer. 2022. "Belief Overreaction and Stock Market Puzzles." Working paper.

Bordalo, Pedro, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer. 2021. "Diagnostic bubbles." *Journal of Financial Economics*, 141(3): 1060–1077.

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar. 2019. "Sticky expectations and the profitability anomaly." *The Journal of Finance*, 74(2): 639–674.

Brandon, Rajna Gibson, and Songtao Wang. 2020. "Earnings belief risk and the cross-section of stock returns." *Review of Finance*, 24(5): 1107–1158.

Brav, Alon, and Reuven Lehavy. 2003. "An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics." *The Journal of Finance*, 58(5): 1933–1967.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich. 2014. "Regression Discontinuity and the Price Effects of Stock Market Indexing." *The Review of Financial Studies*, 28(1): 212–246.

Chen, Long, Zhi Da, and Xinlei Zhao. 2013. "What drives stock price movements?" *The Review of Financial Studies*, 26(4): 841–876.

Chevalier, Judith, and Glenn Ellison. 1997. "Risk taking by mutual funds as a response to incentives." *Journal of political economy*, 105(6): 1167–1200.

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho. 2003. "The value spread." *The Journal of Finance*, 58(2): 609–641.

Crane, Alan D, Sébastien Michenaud, and James P Weston. 2016. "The effect of institutional ownership on payout policy: Evidence from index thresholds." *The Review of Financial Studies*, 29(6): 1377–1408.

Davies, Peter Lloyd, and Michael Canes. 1978. "Stock prices and the publication of second-hand information." *Journal of Business*, 43–56.

De Bondt, Werner FM, and Richard H Thaler. 1990. "Do security analysts overreact?" *The American economic review*, 52–57.

De La O, Ricardo, and Sean Myers. 2021. "Subjective cash flow and discount rate expectations." *The Journal of Finance*, 76(3): 1339–1387.

De La O, Ricardo, and Sean Myers. 2023. "Which Subjective Expectations Explain Asset Prices?"

De la O, Ricardo, Xiao Han, and Sean Myers. 2023. "The Cross-section of Subjective Expectations: Understanding Prices and Anomalies."

Derrien, François, Ambrus Kecskés, and David Thesmar. 2013. "Investor horizons and corporate policies." *Journal of Financial and Quantitative Analysis*, 48(6): 1755–1780.

Dessaint, Olivier, Thierry Foucault, Laurent Frésard, and Adrien Matray. 2019. "Noisy stock prices and corporate investment." *The Review of Financial Studies*, 32(7): 2625–2672.

Di Maggio, Marco, Francesco Franzoni, Shimon Kogan, and Ran Xing. 2023. "Avoiding Idiosyncratic Volatility: Flow Sensitivity to Individual Stock Returns." National Bureau of Economic Research.

Dubey, Pradeep, John Geanakoplos, and Martin Shubik. 1987. "The revelation of information in strategic market games: A critique of rational expectations equilibrium." *Journal of Mathematical Economics*, 16(2): 105–137.

Eckbo, B Espen, Tanakorn Makaew, and Karin S Thorburn. 2018. "Are stock-financed takeovers opportunistic?" *Journal of Financial Economics*, 128(3): 443–465.

Edmans, Alex, Itay Goldstein, and Wei Jiang. 2012. "The real effects of financial markets: The impact of prices on takeovers." *The Journal of Finance*, 67(3): 933–971.

Edmans, Alex, Itay Goldstein, and Wei Jiang. 2015. "Feedback effects, asymmetric trading, and the limits to arbitrage." *American Economic Review*, 105(12): 3766–3797.

Fang, Lily H, and Ayako Yasuda. 2014. "Are stars' opinions worth more? The relation between analyst reputation and recommendation values." *Journal of Financial Services Research*, 46(3): 235–269.

Francis, Jennifer, and Leonard Soffer. 1997. "The relative informativeness of analysts' stock recommendations and earnings forecast revisions." *Journal of Accounting Research*, 35(2): 193–211.

Frankel, Richard, and Charles MC Lee. 1998. "Accounting valuation, market expectation, and cross-sectional stock returns." *Journal of Accounting and economics*, 25(3): 283–319.

Frazzini, Andrea, and Owen A Lamont. 2008. "Dumb money: Mutual fund flows and the cross-section of stock returns." *Journal of financial economics*, 88(2): 299–322.

Funk, Simon. 2006. "Netflix update: Try this at home."

Giammarino, Ronald, Robert Heinkel, Burton Hollifield, and Kai Li. 2004. "Corporate decisions, information and prices: Do managers move prices or do prices move managers?" *Economic Notes*, 33(1): 83–110.

Glaeser, Edward L, and Charles G Nathanson. 2017. "An extrapolative model of house price dynamics." *Journal of Financial Economics*, 126(1): 147–170.

Glossner, Simon. 2019. "Russell index reconstitutions, institutional investors, and corporate social responsibility." *Available at SSRN 3180776.*

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review*, 110(8): 2586–2624.

Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of returns and expected returns." *The Review of Financial Studies*, 27(3): 714–746.

Grossman, Sanford J, and Joseph E Stiglitz. 1980. "On the impossibility of informationally efficient markets." *The American economic review*, 70(3): 393–408.

Groth, John C, Wilbur G Lewellen, Gary G Schlarbaum, and Ronald C Lease. 1979. "An analysis of brokerage house securities recommendations." *Financial Analysts Journal*, 35(1): 32–40.

Heath, Davidson, Daniele Macciocchi, Roni Michaely, and Matthew C Ringgenberg. 2022. "Do index funds monitor?" *The Review of Financial Studies*, 35(1): 91–131.

Hellwig, Martin F. 1980. "On the aggregation of information in competitive markets." *Journal of economic theory*, 22(3): 477–498.

Hribar, Paul, and John McInnis. 2012. "Investor sentiment and analysts' earnings forecast errors." *Management science*, 58(2): 293–307.

Ippolito, Richard A. 1992. "Consumer reaction to measures of poor quality: Evidence from the mutual fund industry." *The Journal of Law and Economics*, 35(1): 45–70.

Irvine, Paul J. 2003. "The incremental impact of analyst initiation of coverage." *Journal of Corporate Finance*, 9(4): 431–451.

Ishigami, Shohei, and Fumiko Takeda. 2018. "Market reactions to stock rating and target price changes in analyst reports: Evidence from Japan." *Journal of International Financial Markets, Institutions and Money*, 52: 134–151.

Jackson, Matthew O. 1991. "Equilibrium, price formation, and the value of private information." *The Review of Financial Studies*, 4(1): 1–16.

Jin, Lawrence J, and Pengfei Sui. 2022. "Asset pricing with return extrapolation." *Journal of Financial Economics*, 145(2): 273–295.

Kerl, Alexander G, and Andreas Walter. 2008. "Never judge a book by its cover: What security analysts have to say beyond recommendations." *Financial Markets and Portfolio Management*, 22(4): 289–321.

Khan, Mozaffar, Leonid Kogan, and George Serafeim. 2012. "Mutual fund trading pressure: Firm-level stock price impact and timing of SEOs." *The Journal of Finance*, 67(4): 1371–1395.

Koijen, Ralph SJ, and Motohiro Yogo. 2019. "A demand system approach to asset pricing." *Journal of Political Economy*, 127(4): 1475–1515.

Kyle, Albert S. 1989. "Informed speculation with imperfect competition." *The Review of Economic Studies*, 56(3): 317–355.

Landier, Augustin, and David Thesmar. 2020. "Earnings expectations during the COVID-19 crisis." *The Review of Asset Pricing Studies*, 10(4): 598–617.

La Porta, Rafael. 1996. "Expectations and the cross-section of stock returns." *The Journal of Finance*, 51(5): 1715–1742.

Lee, Charles MC, and Bhaskaran Swaminathan. 2000. "Price momentum and trading volume." *the Journal of Finance*, 55(5): 2017–2069.

Lee, Charles MC, and Eric C So. 2017. "Uncovering expected returns: Information in analyst coverage proxies." *Journal of Financial Economics*, 124(2): 331–348.

Lee, Charles MC, James Myers, and Bhaskaran Swaminathan. 1999. "What is the Intrinsic Value of the Dow?" *The Journal of Finance*, 54(5): 1693–1741.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter. 2022. "Valid t-ratio Inference for IV." *American Economic Review*, 112(10): 3260–3290.

Li, Jiacui. 2021. "What drives the size and value factors?" *Available at SSRN 2909960.*

Li, Jian, Zhiyu Fu, and Manav Chaudhary. 2022. "Corporate Bond Elasticities: Substitutes Matter." *Available at SSRN 4292266.*

Lou, Dong. 2012. "A flow-based explanation for return predictability." *The Review of Financial Studies*, 25(12): 3457–3489.

Lou, Xiaoxia, and Albert Y Wang. 2018. "Flow-induced trading pressure and corporate investment." *Journal of Financial and Quantitative Analysis*, 53(1): 171–201.

Malkiel, Burton G. 1970. "The valuation of public utility equities." *The Bell Journal of Economics and Management Science*, 143–160.

Mendel, Brock, and Andrei Shleifer. 2012. "Chasing noise." *Journal of Financial Economics*, 104(2): 303–320.

Nagel, Stefan, and Zhengyang Xu. 2021. "Asset Pricing with Fading Memory." *The Review of Financial Studies.*

Norli, Øyvind, Charlotte Ostergaard, and Ibolya Schindele. 2015. "Liquidity and shareholder activism." *The Review of Financial Studies*, 28(2): 486–520.

Office, U.S. Government Publishing. 2023. "17 CFR Â§ 270.30e-1 - Delivery of prospectus by certain open-end management investment companies."

Pancost, N Aaron, and Garrett Schaller. 2022. "Measuring Measurement Error."

Park, Chul W, and Earl K Stice. 2000. "Analyst forecasting ability and the stock price reaction to forecast revisions." *Review of Accounting Studies*, 5(3): 259–272.

Pavlova, Anna, and Taisiya Sikorskaya. 2023. "Benchmarking intensity." *The Review of Financial Studies*, 36(3): 859–903.

Phillips, Gordon M, and Alexei Zhdanov. 2013. "R&D and the incentives from merger and acquisition activity." *The Review of Financial Studies*, 26(1): 34–78.

Schmidt, Cornelius, and Rüdiger Fahlenbrach. 2017. "Do exogenous changes in passive institutional ownership affect corporate governance and firm value?" *Journal of financial economics*, 124(2): 285–306.

Securities, U.S., and Exchange Commission. 2023*a*. "Form N-CSR."

Securities, U.S., and Exchange Commission. 2023*b*. "Form N-Q."

Sharma, Varun. 2023. "The product market effects of index inclusion." *Available at SSRN 3897160*.

Sharpe, Steven A. 2005. "How does the market interpret analysts' long-term growth forecasts?" *Journal of Accounting, Auditing & Finance*, 20(2): 147–166.

Sirri, Erik R, and Peter Tufano. 1998. "Costly search and mutual fund flows." *The journal of finance*, 53(5): 1589–1622.

Stickel, Scott E. 1995. "The anatomy of the performance of buy and sell recommendations." *Financial Analysts Journal*, 51(5): 25–39.

Subrahmanyam, Avanidhar, and Sheridan Titman. 2001. "Feedback from stock prices to cash flows." *The Journal of Finance*, 56(6): 2389–2413.

Wardlaw, Malcolm. 2020. "Measuring mutual fund flow pressure as shock to stock returns." *The Journal of Finance*, 75(6): 3221–3243.

Wei, Wei, and Alex Young. 2021. "Selection bias or treatment effect? A re-examination of Russell 1000/2000 Index reconstitution." *Critical Finance Review*, 1000.

Wharton Research Data Services. 2008. "TF Estimates Glossary." Accessed on March 26, 2023.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data.* MIT press.

Zhu, Qifei. 2020. "The missing new funds." *Management Science*, 66(3): 1193–1204.

Zuo, Luo. 2016. "The informational feedback effect of stock prices on management forecasts." *Journal of Accounting and Economics*, 61(2-3): 391–413.

Internet Appendix

Aditya Chaudhry

January 12, 2024

A Proofs for Section 2

A.1 Proofs for Model from Section 2.2.1

The investor chooses quantity demanded to maximize expected utility under his subjective beliefs (which are the same as those of the analyst, hence the A superscripts in $\mathbb{E}_t^A$ and $\mathbb{V}_t^A$):

$$Q_1^I = \text{argmax}_Q \mathbb{E}_1^A \left[-\exp\left[-A\left(W_1 + Q\left(\tilde{D}_2 - \tilde{P}_1 \right) \right) \right] \right] = \frac{1}{A} \frac{\mathbb{E}_1^A\left[\tilde{D}_2\right] - \tilde{P}_1}{\mathbb{V}_1^A\left[\tilde{D}_2\right]}$$

$$\mathbb{V}_1^A\left[\tilde{D}_2\right] = \sigma_\varepsilon^2.$$

By market clearing, the quantity demanded equals the total number of shares: $Q_1^I = 1$. The market clearing price is:

$$\begin{aligned} \tilde{P}_1 &= \mathbb{E}_1^A\left[\tilde{D}_2\right] - A\mathbb{V}_1^A\left[\tilde{D}_2\right] \\ &= \bar{D} + \tilde{s} - A\sigma_\varepsilon^2. \end{aligned} \tag{24}$$

Proof of Proposition 1. Prices correlate positively with analyst cash flow expectations:

$$Cov\left(\tilde{P}_1, \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) = \sigma_s^2 > 0.$$

The period-one analyst expectation and price negatively predict the period-two return:

$$\text{Period-Two Return} = \tilde{D}_2 - P_1 = -\tilde{s} + A\sigma_\varepsilon^2 + \tilde{\varepsilon}$$

$$Cov\left(\mathbb{E}^A\left[\tilde{D}_2\right], \tilde{D}_2 - P_1\right) = Cov\left(P_1, \tilde{D}_2 - P_1\right) = -\sigma_s^2 < 0.$$

The period-one analyst expectation and price negatively predict the period-two forecast error:

$$\text{Period-Two Forecast Error} = \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right] = -\tilde{s} + \tilde{\varepsilon}$$
$$Cov\left(\mathbb{E}_1^A\left[\tilde{D}_2\right], \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) = Cov\left(P_1, \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) = -\sigma_s^2 < 0.$$

□

A.2 Proofs for Model from Section 2.2.2

The quantity demanded by the CARA investor is

$$Q_1^I = \text{argmax}_Q \mathbb{E}_1^I\left[-\exp\left[-A\left(W_1 + Q\left(\tilde{D}_2 - \tilde{P}_1\right)\right)\right]\right] = \frac{1}{A}\frac{\mathbb{E}_1^I\left[\tilde{D}_2\right] - \tilde{P}_1}{\mathbb{V}_1^I\left[\tilde{D}_2\right]}$$
$$\mathbb{V}_1^I\left[\tilde{D}_2\right] = \sigma_\varepsilon^2.$$

The quantity demanded by the noise trader is:

$$Q_1^N = \tilde{z} \sim N\left(0, \sigma_z^2\right).$$

Market clearing $(1 = Q_1^I + Q_1^N)$ implies the following price:

$$\tilde{P}_1 = \mathbb{E}_1^I\left[\tilde{D}_2\right] - (1 - \tilde{z})\, A\mathbb{V}_1^I\left[\tilde{D}_2\right] \tag{25}$$
$$= \bar{D} + \tilde{\eta} - (1 - \tilde{z})\, A\sigma_\varepsilon^2. \tag{26}$$

The analyst posterior expectation is:

$$\mathbb{E}_1^A\left[\tilde{D}_2\right] = \bar{D} + \frac{\hat{\omega}}{\tau + \hat{\omega}}\left(\tilde{P}_1 - \bar{D} + A\sigma_\varepsilon^2\right)$$
$$= \bar{D} + \frac{\hat{\omega}}{\tau + \hat{\omega}}\tilde{\eta} + \frac{\hat{\omega}}{\tau + \hat{\omega}}A\sigma_\varepsilon^2\tilde{z}. \tag{27}$$

Proof of Proposition 2. Prices correlate positively with analyst cash flow expectations:

$$Cov\left(\tilde{P}_1, \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) = \frac{1}{\tau}\frac{\hat{\omega} + \tau\hat{\omega}\omega^{-1}}{\hat{\omega} + \tau} > 0,$$

where $\omega^{-1} = A^2\sigma_\varepsilon^4\sigma_z^2$. Note that the analyst expectation covaries more strongly with price than the CARA investor's expectation does:

$$Cov\left(\tilde{P}_1, \mathbb{E}_1^I\left[\tilde{D}_2\right]\right) = \frac{1}{\tau} < \frac{1}{\tau}\frac{\hat{\omega} + \tau\hat{\omega}\omega^{-1}}{\hat{\omega} + \tau},$$

since $\hat{\omega}^{-1} < \omega^{-1}$ as $\hat{\sigma}_z < \sigma_z$. If the analyst has the right belief about the distribution of the noise shock $\tilde{z}$ (i.e. $\hat{\sigma}_z = \sigma_z$ so $\hat{\omega} = \omega$), then both of these covariances are equal.

The period-one analyst expectation and price negatively predict the period-two return:

$$\begin{aligned}
\text{Period-Two Return} = \tilde{D}_2 - \tilde{P}_1 &= -A\sigma_\varepsilon^2\tilde{z} + A\sigma_\varepsilon^2 + \tilde{\varepsilon} \\
Cov\left(\mathbb{E}^A\left[\tilde{D}_2\right], \tilde{D}_2 - \tilde{P}_1\right) &= -\frac{\hat{\omega}}{\tau + \hat{\omega}} \cdot \frac{1}{\omega} < 0 \\
Cov\left(\tilde{P}_1, \tilde{D}_2 - \tilde{P}_1\right) &= -\frac{1}{\omega} < 0.
\end{aligned}$$

Note that both of these statements hold even if the analyst has the right belief about the distribution of the noise shock $\tilde{z}$ (i.e. $\hat{\sigma}_z = \sigma_z$ so $\hat{\omega} = \omega$).

The period-one analyst expectation and price negatively predict the period-two forecast error:

$$\begin{aligned}
\text{Period-Two Forecast Error} = \tilde{D}_2 - \mathbb{E}^A\left[\tilde{D}_2\right] &= \frac{\tau}{\tau + \hat{\omega}}\tilde{\eta} - \frac{\hat{\omega}}{\tau + \hat{\omega}}A\sigma_\varepsilon^2\tilde{z} + \tilde{\varepsilon} \\
Cov\left(\mathbb{E}_1^A\left[\tilde{D}_2\right], \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) &= \frac{\hat{\omega}}{(\tau + \hat{\omega})^2}\left(1 - \hat{\omega}\omega^{-1}\right) < 0 \\
Cov\left(\tilde{P}_1, \tilde{D}_2 - \mathbb{E}_1^A\left[\tilde{D}_2\right]\right) &= \frac{1}{\tau + \hat{\omega}}\left(1 - \hat{\omega}\omega^{-1}\right) < 0,
\end{aligned}$$

where both inequalities follow from $\hat{\omega}^{-1} < \omega^{-1}$. If the analyst has the right belief about the distribution of the noise shock $\tilde{z}$ (i.e. $\hat{\sigma}_z = \sigma_z$ so $\hat{\omega} = \omega$ and the analyst does not overreact to price), then both of these covariances equal zero because the analyst would be perfectly-Bayesian and use all public information optimally. Note that the investor forecast error is just $\tilde{\varepsilon}$, which is unpredictable. $\square$

A.3 Proofs for Model from Section 2.2.3

The investor with CARA utility has quantities demanded in periods one and two of:

$$Q_1^I = \text{argmax}_Q \mathbb{E}_1^A \left[-\exp\left[-A \left(W_1 + Q \left(\tilde{P}_2 - \tilde{P}_1 \right) \right) \right] \right] = \frac{1}{A} \frac{\mathbb{E}_1^A \left[\tilde{P}_2 \right] - \tilde{P}_1}{\mathbb{V}_1^A \left[\tilde{P}_2 \right]}$$

$$Q_2^I = \text{argmax}_Q \mathbb{E}_2^A \left[-\exp\left[-A \left(W_2 + Q \left(\tilde{D}_3 - \tilde{P}_2 \right) \right) \right] \right] = \frac{1}{A} \frac{\mathbb{E}_2^A \left[\tilde{D}_3 \right] - \tilde{P}_2}{\mathbb{V}_2^A \left[\tilde{D}_3 \right]}$$

$$\mathbb{V}_2^A \left[\tilde{D}_3 \right] = \sigma_\varepsilon^2.$$

There is an analyst with the same subjective beliefs as the CARA investor (hence the A superscripts in $\mathbb{E}_t^A$ and $\mathbb{V}_t^A$).

The noise trader has demand in periods $t = 1$ and 2 of:

$$Q_t^N = \tilde{z}_t \sim N\left(0, \sigma_z^2\right).$$

Market clearing in each period ($1 = Q_t^I + Q_t^N$) implies the following prices:

$$\tilde{P}_2 = \mathbb{E}_2^A \left[\tilde{D}_3 \right] - (1 - \tilde{z}_2) A \mathbb{V}_2^A \left[\tilde{D}_3 \right] \tag{28}$$

$$= \bar{D} - (1 - \tilde{z}_2) A \sigma_\varepsilon^2 \tag{29}$$

$$\tilde{P}_1 = \mathbb{E}_1^A \left[\tilde{P}_2 \right] - (1 - \tilde{z}_1) A \mathbb{V}_1^A \left[\tilde{P}_2 \right] \tag{30}$$

$$= \frac{1}{1 - \phi} \left[\bar{D} - A\sigma_\varepsilon^2 - (1 - \tilde{z}_1) A^3 \sigma_\varepsilon^4 \sigma_z^2 \right], \tag{31}$$

where (31) follows by substituting (7) into (30) and rearranging.

The CARA investor forms cash flow expectations in period one to rationalize his biased price expectations. That is, this investor has expectations in *period one* about $\tilde{D}_3$ that are consistent with both the law of iterated expectations *and* his biased expectation about the period-two price. From (28) and the law of iterated expectations, we have the period-one expectation of $\tilde{P}_2$ in terms of the period-one expectation of

$\tilde{D}_3$:

$$\mathbb{E}_1^A\left[\tilde{P}_2\right] = \mathbb{E}_1^A\left[\mathbb{E}_2^A\left[\tilde{D}_3\right] - (1-\tilde{z}_2)\, A\mathbb{V}_2^A\left[\tilde{D}_3\right]\right]$$
$$= \mathbb{E}_1^A\left[\tilde{D}_3\right] - A\sigma_\varepsilon^2.$$

Setting this expression equal to (7) and plugging in (29) and (31) yields

$$\begin{aligned}\mathbb{E}_1^A\left[\tilde{D}_3\right] - A\sigma_\varepsilon^2 &= \bar{D} - A\sigma_\varepsilon^2 + \phi\tilde{P}_1 \\ \leftrightarrow \mathbb{E}_1^A\left[\tilde{D}_3\right] &= \bar{D} + \phi\tilde{P}_1 \\ &= \frac{1}{1-\phi}\bar{D} - \frac{\phi}{1-\phi}\left[A\sigma_\varepsilon^2 + (1-\tilde{z}_1)\, A^3\sigma_\varepsilon^4\sigma_z^2\right].\end{aligned} \tag{32}$$

Note that without extrapolation ($\phi = 0$), this investor has rational expectations about $\tilde{D}_3$ in period one: $\mathbb{E}_1^A\left[\tilde{D}_3\right] = \bar{D}$. However, with extrapolation ($\phi > 0$), the period-one noise trader demand shock distorts cash flow expectations.

Proof of Proposition 3. Prices correlate positively with analyst cash flow expectations:

$$Cov\left(\tilde{P}_1, \mathbb{E}_1^A\left[\tilde{D}_3\right]\right) = \frac{\phi}{(1-\phi)^2}\left(A^3\sigma_\varepsilon^4\sigma_z^3\right)^2 > 0.$$

The period-one analyst expectation and price negatively predict the period-two return:

$$\text{Period-Two Return} = \tilde{P}_2 - \tilde{P}_1 = \frac{-\phi}{1-\phi}\left[\bar{D} - A\sigma_\varepsilon^2\right] + \left[\tilde{z}_2 A\sigma_\varepsilon^2 + \frac{1}{1-\phi}(1-\tilde{z}_1)\, A^3\sigma_\varepsilon^4\sigma_z^2\right]$$
$$Cov\left(\mathbb{E}_1^A\left[\tilde{D}_3\right], \tilde{P}_2 - \tilde{P}_1\right) = -\frac{\phi}{(1-\phi)^2}\left(A^3\sigma_\varepsilon^4\sigma_z^3\right)^2 < 0$$
$$Cov\left(\tilde{P}_1, \tilde{P}_2 - \tilde{P}_1\right) = -\frac{1}{(1-\phi)^2}\left(A^3\sigma_\varepsilon^4\sigma_z^3\right)^2 < 0.$$

The period-one analyst expectation and price negatively predict the period-three forecast error:

$$\text{Period-Three Forecast Error} = \tilde{D}_3 - \mathbb{E}_1^A\left[\tilde{D}_3\right] = \frac{-\phi}{1-\phi}\bar{D} + \frac{\phi}{1-\phi}\left[A\sigma_\varepsilon^2 + (1-\tilde{z}_1)\, A^3\sigma_\varepsilon^4\sigma_z^2\right] + \tilde{\varepsilon}$$
$$Cov\left(\mathbb{E}_1^A\left[\tilde{D}_3\right], \tilde{D}_3 - \mathbb{E}_1^A\left[\tilde{D}_3\right]\right) = -\frac{\phi^2}{(1-\phi)^2}\left(A^3\sigma_\varepsilon^4\sigma_z^3\right)^2 < 0$$
$$Cov\left(\tilde{P}_1, \tilde{D}_3 - \mathbb{E}_1^A\left[\tilde{D}_3\right]\right) = -\frac{\phi}{(1-\phi)^2}\left(A^3\sigma_\varepsilon^4\sigma_z^3\right)^2 < 0.$$

□

B Supplements to Section 4

B.1 Description of Russell Banding Methodology Starting in 2007

Prior to 2007, firms with market capitalizations on the May rank date that fell between ranks 1 and 1000 were assigned to the Russell 1000, and those with market caps ranked between 1001 and 3000 were assigned to the Russell 2000.

To reduce turnover, since 2007 FTSE Russell has used a "banding policy" under which there are two separate cutoffs for stocks in the Russell 1000 and 2000 in the previous year, both of which are based on a mechanical function of the firm size distribution in the year. Under this policy:

- Stocks in the Russell 2000 in the previous year are assigned to the Russell 1000 if they're rank date market cap ranks fall between 1 and $1000 - c_1$.
- Stocks in the Russell 1000 in the previous year are assigned to the Russell 2000 if they're rank date market cap ranks fall between $1000 + c_2$ and 3000.

To calculate c_1 and c_2 Russell first computes the cumulative market cap of the largest 1000 stocks (i.e. those with ranks 1 through 1000). Let $C(N)$ represent the cumulative market cap of the largest N stocks. c_1 is calculated such that $C(1000 - c_1) = 0.95 \cdot C(1000)$. c_2 is calculated such that $C(1000 + c_2) = 1.05 \cdot C(1000)$. That is, the band of stocks between ranks $1000 - c_1$ and $1000 + c_2$ constitutes a 5% band around the cumulative market cap of the largest 1000 stocks.

Thus, even after the introduction of the banding policy, assignment to the Russell 1000 or 2000 is still based on a mechanical rule. After the introduction of the banding policy, this mechanical rule changes each year with the distribution of firm sizes.

B.2 Supplemental Empirical Results

Table B1: Predicting Russell 2000 Index Membership in June

	(1) BW = 100	(2) BW = 125	(3) BW = 150	(4) BW = 175	(5) BW = 200
1(Rank > Cutoff in May)	0.754***	0.761***	0.756***	0.753***	0.744***
	(0.0543)	(0.0547)	(0.0575)	(0.0575)	(0.0592)
$\log$ Mkt $\text{Cap}_{n,t}$	-0.199***	-0.205***	-0.219***	-0.232***	-0.244***
	(0.0567)	(0.0590)	(0.0620)	(0.0570)	(0.0549)
$1(\text{In Band in May})_{n,t}$	-0.0342**	-0.0382**	-0.0523**	-0.0644***	-0.0726***
	(0.0137)	(0.0173)	(0.0207)	(0.0211)	(0.0229)
$1(\text{In Rusell 2000 in May})_{n,t}$	0.0843**	0.0703**	0.0609*	0.0492	0.0401
	(0.0337)	(0.0335)	(0.0339)	(0.0301)	(0.0284)
$1(\text{In Band in May})_{n,t} \cdot 1(\text{In Rusell 2000 in May})_{n,t}$	0.0762**	0.0846**	0.105**	0.124***	0.140***
	(0.0290)	(0.0353)	(0.0412)	(0.0427)	(0.0456)
Bid-Ask % $\text{Spread}_{n,t}$	0.525	0.684	0.658	0.856	0.742
	(0.634)	(0.592)	(0.600)	(0.598)	(0.535)
Year FE	Y	Y	Y	Y	Y
Year-Clustered SE	Y	Y	Y	Y	Y
N	5773	7227	8649	10079	11251
F	35291.8	25529.5	20118.9	16585.3	13904.0
R-Squared	0.849	0.863	0.874	0.883	0.888

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

This table displays results for the following regression:

$$1(\text{Stock } n \in \text{Rusell 2000 in June}) = b_0 + b_1 1(\text{Stock } n \text{ Rank Date Mkt Cap} > \text{Cutoff}) + \boldsymbol{\beta}_1' \boldsymbol{X}_{n,t} + FE_t + e_{1,n,t}.$$

$\boldsymbol{X}_{n,t}$ includes the log market cap as of the May rank date, the one-year monthly rolling average bid-ask percentage spread, and the banding controls (an indicator for having rank-date market cap in the band including stocks from the Russell 1000 and 2000, an indicator for being in the Russell 2000 in May before reconstitution, and the interaction of these indicators). Columns indicate the bandwidth around the cutoff used. This regression only includes (stock, year) observations for which analyst expectations are available. The time period is 1999-05:2018-09.

Market caps are calculated using the methodology of Ben-David, Franzoni and Moussawi (2019). At a high level, this methodology computes market cap as the maximum of: CRSP market cap, COMPUSTAT market cap for only tradeable share classes, and COMPUSTAT market cap for all share classes. I use the SAS code provided by the authors verbatim, except for calculating the market cap of Berkshire Hathaway for which (due to abnormalities in the COMPUSTAT market cap for all share classes) I use the maximum of the CRSP market cap and COMPUSTAT market cap for only tradeable share classes.

Table B2: Full Results for Effect of Prices on Annual EPS Expectations Using ΔBMI as Instrument

	(1) OLS	(2) OLS	(3) OLS	(4) First Stage	(5) First Stage	(6) First Stage	(7) Reduced Form	(8) Reduced Form	(9) Reduced Form	(10) 2SLS	(11) 2SLS	(12) 2SLS
$\Delta p_{a,n,t}$	0.245***	0.265***	0.265***							0.209	0.416**	0.439***
	(0.0123)	(0.0111)	(0.00984)							(-0.320, 0.738)	(0.0639, 0.768)	(0.215, 0.663)
$\Delta BMI_{n,t}$				0.544***	0.573***	0.575***	0.114	0.238***	0.252***			
				(0.162)	(0.168)	(0.119)	(0.103)	(0.0736)	(0.0759)			
log Mkt Cap$_{n,t}$	0.0300***	0.0539***	0.0478***	0.0302	0.0969***	0.0366**	0.0369**	0.0846***	0.0637***	0.0306***	0.0442***	0.0476***
	(0.00865)	(0.00790)	(0.00939)	(0.0639)	(0.0157)	(0.0170)	(0.0158)	(0.0105)	(0.0134)	(0.00644)	(0.00858)	(0.00849)
$D_{1,n,t} = 1(\text{In Band in May})_{n,t}$	-0.000902	0.00597	0.00112	0.0130	0.0227**	0.0225**	0.00203	0.0138*	0.00926*	-0.000694	0.00440	-0.000625
	(0.00599)	(0.00495)	(0.00336)	(0.0156)	(0.0105)	(0.00943)	(0.00832)	(0.00668)	(0.00447)	(0.00580)	(0.00433)	(0.00350)
$D_{2,n,t} = 1(\text{In Rusell 2000 in May})_{n,t}$	0.0430***	0.0507***	0.0520***	0.0705**	0.0989***	0.0873***	0.0596***	0.0808***	0.0803***	0.0449***	0.0397***	0.0420***
	(0.00817)	(0.00826)	(0.00586)	(0.0326)	(0.0142)	(0.0138)	(0.0107)	(0.0102)	(0.00802)	(0.00744)	(0.0102)	(0.00783)
$D_{1,n,t} \cdot D_{2,n,t}$	-0.00311	-0.0105	0.000294	-0.00590	-0.0350*	-0.0224	-0.00384	-0.0238**	-0.0107	-0.00260	-0.00919	-0.000875
	(0.00744)	(0.00722)	(0.00523)	(0.0221)	(0.0179)	(0.0170)	(0.0115)	(0.00990)	(0.00845)	(0.00898)	(0.00694)	(0.00549)
Bid-Ask % Spread$_{n,t}$	0.159	0.749	2.747***	2.535	5.620***	4.264***	0.769	2.233***	3.896***	0.238	-0.105	2.024**
	(0.776)	(0.539)	(0.775)	(2.581)	(1.022)	(1.422)	(1.018)	(0.642)	(0.989)	(0.843)	(0.550)	(0.858)
Year FE		Y	Y		Y	Y		Y	Y		Y	Y
Stock FE			Y			Y			Y			Y
Year-Clustered SE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	164512	164512	164492	164512	164512	164492	164512	164512	164492	164512	164512	164492
(Partial) F	92.95	109.0	235.5	10.72	11.68	23.33	15.71	18.27	26.27	18.67	21.67	38.51
R-Squared	0.0990	0.103	0.0844	0.0213	0.0319	0.0175	0.0175	0.0234	0.0151			

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

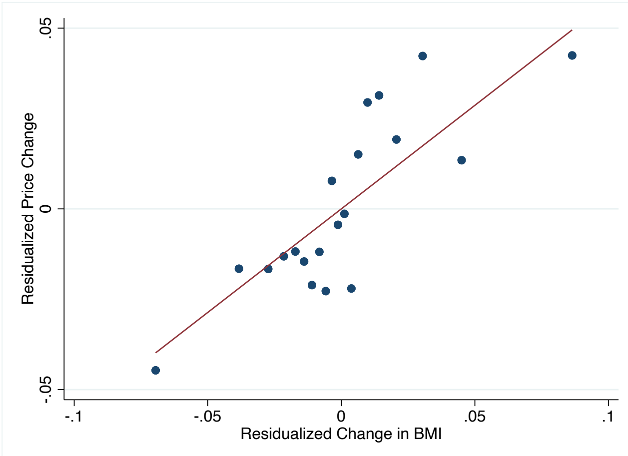
This table displays results for the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \Delta BMI_{n,t} + \boldsymbol{\beta}_1' \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$$

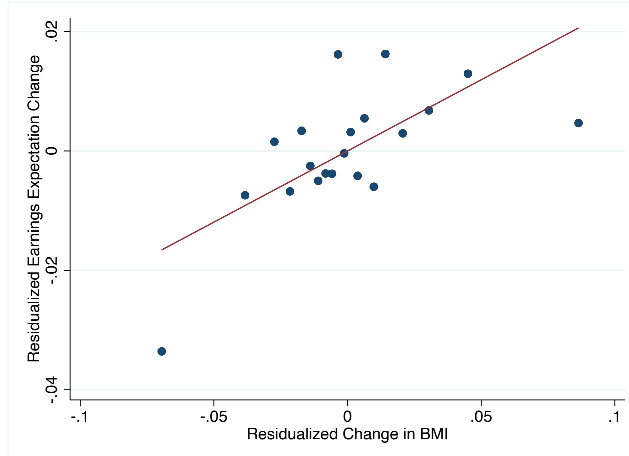
$$\Delta E_{a,n,t+4h|t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{\beta}_2' \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t,h},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the June change in BMI ($\Delta BMI_{n,t}$). The reduced form regresses quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on $\Delta BMI_{n,t}$. The second stage regresses $\Delta E_{a,n,t+4h|t}$ on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). $\boldsymbol{X}_{n,t}$ includes the log market cap as of the May rank date, the one-year monthly rolling average bid-ask percentage spread, and the banding controls (an indicator for having rank-date market cap in the band including stocks from the Russell 1000 and 2000, an indicator for being in the Russell 2000 in May before reconstitution, and the interaction of these indicators). Columns 4 through 7 report the partial F statistics associated with the first-stage regressions; the other columns report the raw F statistics. Columns 10 through 12 report the 95% confidence interval for the second-stage coefficient using the tF procedure of Lee et al. (2022). All units are in percentage points (i.e. 1.0 is 1%). The time period is 1999-05:2018-09.

Figure B1: Binscatter for First Stage and Reduced Form for ΔBMI Instrument



(a) First-Stage



(b) Reduced-Form

This figure displays binscatter plots for the following first-stage and reduced-form regressions:

$$\Delta p_{a,n,t} = a_0 + a_1 \Delta BMI_{n,t} + \boldsymbol{\beta}_1' \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta E_{a,n,t+4h|t} = b_0 + b_1 \Delta BMI_{n,t} + \boldsymbol{\beta}_2' \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t,h},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the June change in BMI ($\Delta BMI_{n,t}$). The reduced form regresses quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on $\Delta BMI_{n,t}$. $\boldsymbol{X}_{n,t}$ includes the log market cap as of the May rank date, the one-year monthly rolling average bid-ask percentage spread, and the banding controls (an indicator for having rank-date market cap in the band including stocks from the Russell 1000 and 2000, an indicator for being in the Russell 2000 in May before reconstitution, and the interaction of these indicators). All units are in percentage points (i.e. 1.0 is 1%). The time period is 1999-05:2018-09.

Table B3: Horizon-Specific Effects of Prices on Cash Flow Expectations Using ΔBMI as Instrument

	OLS	First Stage	Reduced Form	2SLS
Panel (a): LTG Expectations				
$\Delta p_{a,n,t}$	0.0580***			0.0210
	(0.00637)			(-0.227, 0.269)
$\Delta BMI_{n,t}$		0.389**	0.00817	
		(0.142)	(0.0250)	
N	4489	4489	4489	4489
F	82.77	7.487	0.107	0.124
R-Squared	0.0646	0.00416	0.0000352	
Panel (b): 1 Year EPS Expectations				
$\Delta p_{a,n,t}$	0.279***			0.501**
	(0.0139)			(0.0419, 0.960)
$\Delta BMI_{n,t}$		0.565***	0.283***	
		(0.157)	(0.0810)	
N	79053	79053	79053	79053
F	402.2	13.05	12.22	10.89
R-Squared	0.0820	0.00749	0.00197	
Panel (b): 2 Year EPS Expectations				
$\Delta p_{a,n,t}$	0.264***			0.388**
	(0.00987)			(0.0639, 0.712)
$\Delta BMI_{n,t}$		0.577***	0.224**	
		(0.167)	(0.0829)	
N	64986	64986	64986	64986
F	715.0	11.88	7.280	14.19
R-Squared	0.0928	0.00756	0.00152	
Panel (c): 3 Year EPS Expectations				
$\Delta p_{a,n,t}$	0.212***			0.136
	(0.00976)			(-0.466, 0.738)
$\Delta BMI_{n,t}$		0.536**	0.0728	
		(0.212)	(0.0837)	
N	16715	16715	16715	16715
F	469.8	6.420	0.757	1.158
R-Squared	0.0602	0.00614	0.000152	
Panel (d): 4 Year EPS Expectations				
$\Delta p_{a,n,t}$	0.258***			0.197
	(0.0144)			(., .)
$\Delta BMI_{n,t}$		0.643*	0.127	
		(0.352)	(0.154)	
N	3758	3758	3758	3758
F	321.8	3.330	0.677	1.017
R-Squared	0.0713	0.00754	0.000315	
Quarter FE	Y	Y	Y	Y
Quarter-Clustered SE	Y	Y	Y	Y

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

This table displays results for the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \Delta BMI_{n,t} + \boldsymbol{\beta}_1' \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{\beta}_2' \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the June change in BMI ($\Delta BMI_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \text{LTG}_{a,n,t}$ in Panel (a)) or quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$ in Panels (b) through (d)) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). $\boldsymbol{X}_{n,t}$ includes the log market cap as of the May rank date, the one-year monthly rolling average bid-ask percentage spread, and the banding controls (an indicator for having rank-date market cap in the band including stocks from the Russell 1000 and 2000, an indicator for being in the Russell 2000 in May before reconstitution, and the interaction of these indicators). Column 4 reports the 95% confidence interval for the second-stage coefficient using the tF procedure of Lee et al. (2022). If the first-stage is not significant at the 5% significance level (as for the four-year EPS expectations in Panel (d)), this procedure assigns the whole real line as the confidence interval for the second-stage coefficient. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1999-05:2018-09.

C Shift-Share Instrument Proofs

This appendix formalizes the argument that exogenous mutual fund ownership shares are sufficient for the FIT instrument to be cross-sectionally exogenous. In particular, Proposition 4 proves the consistency of the two-stage least squares estimator for α that uses $\text{FIT}_{n,t}$ as an instrument for price under the assumption of exogenous ownership shares. The arguments in this appendix are the same as those in Goldsmith-Pinkham, Sorkin and Swift (2020); I include them here only for completeness.

In particular, I fix an analyst (so I drop the analyst a subscript). In this case, we have the following system of equations in the cross section of stocks:

$$\Delta p_{n,t} = M\text{FIT}_{n,t} + \boldsymbol{X}'_{n,t}\boldsymbol{\delta}_1 + \epsilon_{n,t}$$
$$\Delta y_{n,t} = \alpha \Delta p_{n,t} + \boldsymbol{X}'_{n,t}\delta_2 + \nu_{n,t}.$$

Empirically I pool across all analysts to increase power.

Appendix C.1 describes the sampling procedure to formalize the idea that the source of variation is the cross section of equities.

Appendix C.2 proves the consistency of the two-stage least squares estimator for α that uses $\text{FIT}_{n,t}$ as an instrument for price under the assumption of exogenous ownership shares.

C.1 Sampling Procedure

Let N be the number of stocks, I be the number of mutual funds, and $T+2$ be the number of quarters. Label the quarters $t = -1, 0, 1, \ldots, T$. Let D be the dimension of $\boldsymbol{X}_{n,t}$ (i.e. the number of controls). Let $L = N \cdot T$.

Define

$$
\begin{aligned}
\boldsymbol{F}_t &= (f_{1,t}, \ldots, f_{I,t}) \in \mathbb{R}^I \\
\boldsymbol{F} &= (F_1, \ldots, F_T) \in \mathbb{R}^{(I \cdot T)\times 1} \\
\boldsymbol{S}_{n,t} &= (S_{1,n,t}, \ldots, S_{I,n,t}) \in \mathbb{R}^I \\
\mathbf{S}_n &= \begin{pmatrix} \boldsymbol{S}^{'}_{n,-1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{S}^{'}_{n,0} & & \\ \vdots & & \ddots & \\ 0 & & & \boldsymbol{S}^{'}_{n,T-2} \end{pmatrix} \in \mathbb{R}^{T\times(I\cdot T)} \\
\mathbf{S} &= (\mathbf{S}_1, \ldots, \mathbf{S}_N) \in \mathbb{R}^{L\times(I\cdot T)} \\
\Delta\boldsymbol{y}_n &= (\Delta y_{n,1}, \ldots, \Delta y_{n,T}) \in \mathbb{R}^T \\
\Delta\mathbf{y} &= (\Delta\boldsymbol{y}_1, \ldots, \Delta\boldsymbol{y}_N) \in \mathbb{R}^L \\
\Delta\boldsymbol{p}_n &= (\Delta p_{n,1}, \ldots, \Delta p_{n,T}) \in \mathbb{R}^T \\
\Delta\mathbf{p} &= (\Delta\boldsymbol{p}_1, \ldots, \Delta\boldsymbol{p}_N) \in \mathbb{R}^L \\
\boldsymbol{\nu}_n &= (\nu_{n,1}, \ldots, \nu_{n,T}) \in \mathbb{R}^T \\
\boldsymbol{\nu} &= (\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N) \in \mathbb{R}^L \\
\mathbf{X}_n &= (\boldsymbol{X}_{n,1}, \ldots, \boldsymbol{X}_{n,T}) \in \mathbb{R}^{D\times T} \\
\mathbf{X} &= \left(\mathbf{X}^{'}_1, \ldots, \mathbf{X}^{'}_N\right) \in \mathbb{R}^{L\times D} \\
\mathbf{FIT}_n &= (\mathrm{FIT}_{n,1}, \ldots, \mathrm{FIT}_{n,T}) \in \mathbb{R}^T \\
\mathbf{FIT} &= (\mathbf{FIT}_1, \ldots, \mathbf{FIT}_N) \in \mathbb{R}^L.
\end{aligned}
$$

Note that

$$
\mathbf{FIT} = \mathbf{S}\boldsymbol{F}. \tag{33}
$$

Let

$$
\mathbf{P}_X = \mathbf{X}\left(\mathbf{X}^{'}\mathbf{X}\right)^{-1}\mathbf{X}^{'},
$$

be the projection matrix with respect to the stacked matrix of controls $\mathbf{X}$ and

$$
\mathbf{M}_X = \mathbf{I}_L - \mathbf{P}_X,
$$

be the annihilator matrix, where $\mathbf{I}_L$ is the $L \times L$ identity matrix.

I assume the data

$$\{\Delta \boldsymbol{y}_n, \Delta \boldsymbol{p}_n, \mathbf{S}_n, \mathbf{X}_n\}_{n=1}^{N}$$

are drawn i.i.d. across stocks n. I do *not* assume the data are i.i.d. within n.

I view the vector of flows $\boldsymbol{F}$ as fixed. That is, flows are viewed as constants, not random variables. This sampling procedure draws i.i.d. samples of $\{\Delta \boldsymbol{y}_n, \Delta \boldsymbol{p}_n, \mathbf{S}_n, \mathbf{X}_n\}$ across stocks n. The flows $f_{i,t}$ are at the fund i and quarter t level and do not vary across stocks n. Thus, I view flows as fixed. This treatment of the flow vector $\boldsymbol{F}$ as fixed is analogous to the treatment of the matrix of national average industry growth rates as fixed in Goldsmith-Pinkham, Sorkin and Swift (2020).

C.2 Proof of Consistency

The estimator for α using the FIT instrument in the following two-stage least squares regression

$$\Delta p_{n,t} = M\text{FIT}_{n,t} + \boldsymbol{X}_{n,t}^{'}\boldsymbol{\delta}_1 + e_{1,n,t}$$
$$\Delta y_{n,t} = \alpha \Delta \hat{p}_{n,t} + \boldsymbol{X}_{n,t}^{'}\delta_2 + e_{2,n,t},$$

is

$$\hat{\alpha}_{\text{FIT}} = \frac{(\mathbf{M}_X \mathbf{FIT})^{'}(\mathbf{M}_X \Delta \mathbf{y})}{(\mathbf{M}_X \mathbf{FIT})^{'}(\mathbf{M}_X \Delta \mathbf{p})} = \frac{\mathbf{FIT}^{'}\mathbf{M}_X \Delta \mathbf{y}}{\mathbf{FIT}^{'}\mathbf{M}_X \Delta \mathbf{p}}.$$

Define a GMM estimator for α that uses the ownership shares $S_{i,n,t-2}$ as cross-sectional instruments for $\Delta p_{n,t}$ (i.e. use $[S_{i,n,s} \cdot 1(s = t-2)]_{i=1,s=-1}^{I,T-2}$ as instruments for $\Delta p_{n,t}$) and weighting matrix $\mathbf{W}$ as

$$\hat{\alpha}_{\text{GMM}} = \frac{\Delta \mathbf{p}^{'}\mathbf{M}_X \mathbf{S}\mathbf{W}\mathbf{S}^{'}\mathbf{M}_X \Delta \mathbf{y}}{\Delta \mathbf{p}^{'}\mathbf{M}_X \mathbf{S}\mathbf{W}\mathbf{S}^{'}\mathbf{M}_X \Delta \mathbf{p}}.$$

The following lemma (which is the same as Proposition 1 from Goldsmith-Pinkham, Sorkin and Swift (2020)) demonstrates that $\hat{\alpha}_{\text{FIT}} = \hat{\alpha}_{\text{GMM}}$ when $\mathbf{W} = \boldsymbol{F}\boldsymbol{F}^{'}$. That is, using the FIT instrument as an instrument for price to estimate α is equivalent to using the ownership shares themselves interacted with quarter indicators as instruments for price in an over identified system with a particular GMM weighting matrix.

Lemma 1 (Equivalence of FIT and GMM Estimators)**.** *$\hat{\alpha}_{FIT} = \hat{\alpha}_{GMM}$ if $\mathbf{W} = \boldsymbol{F}\boldsymbol{F}^{'}$.*

Proof.

$$\begin{aligned}
\hat{\alpha}_{\mathrm{FIT}} &= \frac{\mathbf{FIT}^{'}\mathbf{M}_X\Delta\mathbf{y}}{\mathbf{FIT}^{'}\mathbf{M}_X\Delta\mathbf{p}} \\
&= \frac{\boldsymbol{F}^{'}\mathbf{S}^{'}\mathbf{M}_X\Delta\mathbf{y}}{\boldsymbol{F}^{'}\mathbf{S}^{'}\mathbf{M}_X\Delta\mathbf{p}} \\
&= \frac{\Delta\mathbf{p}^{'}\mathbf{M}_X\mathbf{S}\boldsymbol{F}\boldsymbol{F}^{'}\mathbf{S}^{'}\mathbf{M}_X\Delta\mathbf{y}}{\Delta\mathbf{p}^{'}\mathbf{M}_X\mathbf{S}\boldsymbol{F}\boldsymbol{F}^{'}\mathbf{S}^{'}\mathbf{M}_X\Delta\mathbf{p}} \\
&= \hat{\alpha}_{\mathrm{GMM}}
\end{aligned}$$

The first equality follows by the definition of $\hat{\alpha}_{\mathrm{FIT}}$. The second follows from (33). The third follows since $\Delta\mathbf{p}^{'}\mathbf{M}_X\mathbf{S}\boldsymbol{F}$ is a scalar. The last equality follows by the definition of $\hat{\alpha}_{\mathrm{GMM}}$ when setting $\mathbf{W} = \boldsymbol{F}\boldsymbol{F}^{'}$. □

Thus, in order to prove the consistency of $\hat{\alpha}_{\mathrm{FIT}}$, it suffices to prove the consistency of $\hat{\alpha}_{\mathrm{GMM}}$.

Proving the consistency of $\hat{\alpha}_{\mathrm{GMM}}$ requires the following three assumptions.

Assumption 1. $0 < \mathbb{E}\left[FIT_{n,t}^2\right] < \infty, \forall t.$

Assumption 2 (Relevance)**.** *In*

$$\Delta p_{n,t} = M\,FIT_{n,t} + \boldsymbol{X}_{n,t}^{'}\boldsymbol{\delta}_1 + \epsilon_{n,t}$$

where $\mathbb{E}\left[\epsilon_{n,t} \mid FIT_{n,t}, \boldsymbol{X}_{n,t}\right] = 0$*, we have* $M \neq 0$ *and* $|M| < \infty$*.*

Assumption 3 (Strict Exogeneity)**.** $\mathbb{E}\left[S_{i,n,t-2}\nu_{n,t} \mid \boldsymbol{X}_{n,t}\right] = 0, \forall i, t,$ *such that* $f_{i,t} \neq 0$*.*

The following proposition proves that $\hat{\alpha}_{\mathrm{GMM}}$ is a consistent estimator for α under Assumptions 1, 2, and 3.

Proposition 4 (Consistency)**.** *Given Assumptions 1, 2, and 3,* $plim_{N\to\infty}\hat{\alpha}_{GMM} - \alpha = 0$ *if* $\mathbf{W} = \boldsymbol{F}\boldsymbol{F}^{'}$*.*

Proof. This proposition follows from standard GMM consistency results. In particular, the proof below follows from the proof of Theorem 8.1 in Wooldridge (2010).

Rewrite $\hat{\alpha}_{\mathrm{GMM}}$ as

$$\begin{aligned}
\hat{\alpha}_{\mathrm{GMM}} &= \frac{\Delta\mathbf{p}^{'}\mathbf{M}_X\mathbf{S}\boldsymbol{F}\boldsymbol{F}^{'}\mathbf{S}^{'}\mathbf{M}_X\Delta\mathbf{y}}{\Delta\mathbf{p}^{'}\mathbf{M}_X\mathbf{S}\boldsymbol{F}\boldsymbol{F}^{'}\mathbf{S}^{'}\mathbf{M}_X\Delta\mathbf{p}} \\
&= \alpha + \frac{\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)^{'}\mathbf{W}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\right)}{\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)^{'}\mathbf{W}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)},
\end{aligned} \tag{34}$$

where

$$\Delta\mathbf{p}^{\perp} = \mathbf{M}_X \Delta\mathbf{p} = \left(\Delta\boldsymbol{p}_1^{\perp}, \ldots, \Delta\boldsymbol{p}_N^{\perp}\right)$$
$$\boldsymbol{\nu}^{\perp} = \mathbf{M}_X \boldsymbol{\nu} = \left(\boldsymbol{\nu}_1^{\perp}, \ldots, \boldsymbol{\nu}_N^{\perp}\right).$$

By the law of large numbers and continuous mapping theorem, the denominator of the second term in (34) converges to

$$\begin{aligned}
\text{plim}_{N\to\infty} \left(\frac{1}{N}\sum_{n=1}^{N} \mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right)^{'} \mathbf{W} \left(\frac{1}{N}\sum_{n=1}^{N} \mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right) &= \mathbb{E}\left[\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right]^{'} \mathbf{W} \mathbb{E}\left[\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right] \\
&= \mathbb{E}\left[\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right]^{'} \boldsymbol{F}\boldsymbol{F}^{'} \mathbb{E}\left[\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right] \\
&= \left(\mathbb{E}\left[\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right]^{'} \boldsymbol{F}\right)^2 \\
&= \left(\mathbb{E}\left[\boldsymbol{F}^{'}\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right]^{'}\right)^2 \\
&= \left(\mathbb{E}\left[\mathbf{FIT}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right]^{'}\right)^2 \\
&\neq 0. \qquad (35)
\end{aligned}$$

The second equality follows since $\mathbf{W} = \boldsymbol{F}\boldsymbol{F}^{'}$. The third equality follows since $\boldsymbol{F}^{'}\mathbb{E}\left[\mathbf{S}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right]$ is a scalar. The fourth equality follows since flows are viewed as fixed, as described in the sampling procedure in Appendix C.1. The fifth equality follows from (33). The last line follows from Assumption 2 since

$$0 \neq M = \mathbb{E}\left[\mathbf{FIT}_n^{'}\mathbf{FIT}_n\right]^{-1} \mathbb{E}\left[\mathbf{FIT}_n^{'} \Delta\boldsymbol{p}_n^{\perp}\right],$$

and $\mathbb{E}\left[\mathbf{FIT}_n^{'}\mathbf{FIT}_n\right] \neq 0$ by Assumption 1.

By the law of large numbers and continuous mapping theorem, the numerator of the second term in

(34) converges to

$$\text{plim}_{N\to\infty}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)^{'}\mathbf{W}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\right)=\mathbb{E}\left[\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right]^{'}\mathbf{W}\mathbb{E}\left[\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\right]$$
$$=\mathbb{E}\left[\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right]^{'}\boldsymbol{F}\boldsymbol{F}^{'}\mathbb{E}\left[\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\right]$$
$$=\mathbb{E}\left[\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right]^{'}\boldsymbol{F}\cdot 0$$
$$=0. \tag{36}$$

The third equality follows since

$$\boldsymbol{F}^{'}\mathbb{E}\left[\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\right]=\boldsymbol{F}^{'}\mathbb{E}\left[\mathbb{E}\left[\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\mid\boldsymbol{X}_{n,t}\right]\right] \tag{37}$$
$$=\sum_{i=1}^{I}\sum_{t=1}^{T}f_{i,t}\mathbb{E}\left[\mathbb{E}\left[S_{i,n,t-2}\nu_{n,t}^{\perp}\mid\boldsymbol{X}_{n,t}\right]\right]$$
$$=0. \tag{38}$$

(37) follows by the law of iterated expectations. (38) follows from Assumption 3, under which[48]

$$\mathbb{E}\left[S_{i,n,t-2}\nu_{n,t}\mid\boldsymbol{X}_{n,t}\right]=0,\forall i,t\text{ s.t. }f_{i,t}\neq 0.$$

Thus, by (35), (36), and the continuous mapping theorem, the second term in (34) converges to zero:

$$\text{plim}_{N\to\infty}\frac{\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)^{'}\mathbf{W}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\boldsymbol{\nu}_n^{\perp}\right)}{\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)^{'}\mathbf{W}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_n^{'}\Delta\boldsymbol{p}_n^{\perp}\right)}=0.$$

Therefore, $\text{plim}_{N\to\infty}\hat{\alpha}_{\text{GMM}}-\alpha=0$. □

[48] Note that

$$\mathbb{E}\left[S_{i,n,t-2}\nu_{n,t}^{\perp}\mid\boldsymbol{X}_{n,t}\right]=\mathbb{E}\left[S_{i,n,t-2}\left(\nu_{n,t}-\mathbb{E}\left[\nu_{n,t}\mid\boldsymbol{X}_{n,t}\right]\right)\mid\boldsymbol{X}_{n,t}\right]=\mathbb{E}\left[S_{i,n,t-2}\nu_{n,t}\mid\boldsymbol{X}_{n,t}\right],$$

because $\mathbb{E}\left[\nu_{n,t}\mid\boldsymbol{X}_{n,t}\right]=0,\forall t$, under the assumption that the controls $\boldsymbol{X}_{n,t}$ are exogenous.

D Supplementary Empirical Results for FIT Instrument

Table D4: Summary Statistics for FIT Instrument

	$\Delta LTG_{a,n,t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	121553.00	121553.00	121553.00	131333.00
Mean	-0.01	0.03	-0.0001	165.30
Std Dev.	0.04	0.22	0.0039	309.02
Min	-0.12	-0.94	-0.3941	1.00
25%	-0.02	-0.08	-0.0010	44.00
50%	-0.00	0.02	-0.0000	73.00
75%	0.01	0.13	0.0007	132.00
Max	0.10	5.80	0.1845	3712.00

(a) LTG Expectations

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	3396550.00	3396550.00	3396550.00	133902.00
Mean	-0.01	0.03	-0.0001	162.61
Std Dev.	0.19	0.24	0.0042	306.68
Min	-0.59	-0.99	-0.4536	1.00
25%	-0.08	-0.09	-0.0011	43.00
50%	0.00	0.02	-0.0000	72.00
75%	0.07	0.14	0.0008	130.00
Max	0.44	14.05	0.2134	3712.00

(b) Annual EPS Expectations — All Horizons

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	1791810.00	1791810.00	1791810.00	133783.00
Mean	-0.03	0.03	-0.0000	162.73
Std Dev.	0.19	0.23	0.0041	306.79
Min	-0.55	-0.99	-0.4536	1.00
25%	-0.09	-0.09	-0.0010	43.00
50%	-0.00	0.02	0.0000	72.00
75%	0.07	0.14	0.0008	130.00
Max	0.37	9.99	0.2134	3712.00

(c) Annual EPS Expectations — 1 Year

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	1294912.00	1294912.00	1294912.00	133778.00
Mean	0.00	0.03	-0.0001	162.74
Std Dev.	0.18	0.24	0.0043	306.79
Min	-0.44	-0.99	-0.4536	1.00
25%	-0.07	-0.09	-0.0012	43.00
50%	0.01	0.02	-0.0000	72.00
75%	0.08	0.14	0.0009	130.00
Max	0.40	14.05	0.2028	3712.00

(d) Annual EPS Expectations — 2 Year

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	247140.00	247140.00	247140.00	126829.00
Mean	-0.00	0.02	-0.0003	164.69
Std Dev.	0.17	0.23	0.0040	308.29
Min	-0.45	-0.97	-0.1185	1.00
25%	-0.06	-0.09	-0.0016	44.00
50%	0.00	0.02	-0.0002	73.00
75%	0.06	0.13	0.0009	132.00
Max	0.39	12.94	0.1845	3712.00

(e) Annual EPS Expectations — 3 Year

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\text{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num Obs.	60438.00	60438.00	60438.00	110952.00
Mean	-0.02	0.03	-0.0002	173.56
Std Dev.	0.21	0.25	0.0039	317.66
Min	-0.59	-0.96	-0.0813	1.00
25%	-0.08	-0.09	-0.0015	46.00
50%	-0.00	0.02	-0.0001	76.00
75%	0.06	0.14	0.0009	141.00
Max	0.44	7.21	0.1845	3712.00

(f) Annual EPS Expectations — 4 Year

Summary statistics for quarter-over-quarter changes in LTG expectations $\Delta LTG_{a,n,t}$ (Panel (a)) and revisions in annual EPS expectations for forecast horizons of one to four years $\Delta E_{a,n,t+4h|t}$ (Panels (b) through (d)), inter-announcement percentage price changes ($\Delta p_{a,n,t}$), the FIT instrument $\text{FIT}_{n,t}$, and the number of stocks held by mutual funds used to construct the FIT instrument. The first three columns are expressed in absolute terms (i.e. 0.01 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Table D5: Effect of Prices on Analyst Expectations Using FIT as Instrument — Quarterly Block-Bootstrapped Confidence Intervals

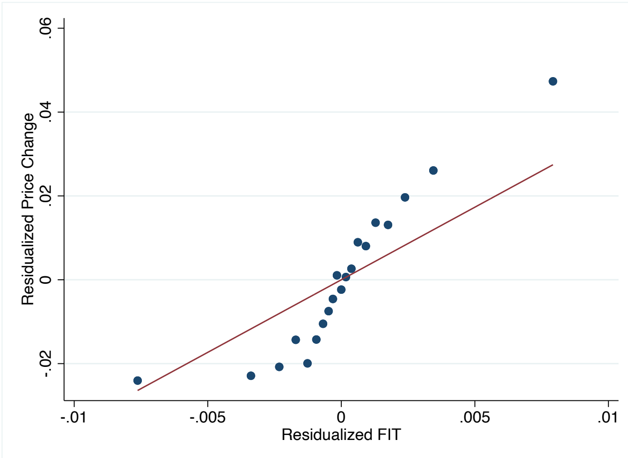
	OLS	First Stage	Reduced Form	2SLS
LTG	0.044	3.576	0.177	0.053
	(0.040, 0.048)	(1.848, 5.546)	(0.088, 0.339)	(0.030, 0.075)
Pooled Annual EPS Expectations	0.282	3.560	0.701	0.214
	(0.259, 0.302)	(2.054, 5.741)	(0.390, 1.276)	(0.115, 0.289)
1-Year EPS Expectations	0.291	3.469	0.709	0.220
	(0.268, 0.313)	(2.048, 5.651)	(0.397, 1.290)	(0.122, 0.315)
2-Year EPS Expectations	0.278	3.588	0.644	0.202
	(0.254, 0.299)	(2.006, 5.866)	(0.357, 1.254)	(0.106, 0.270)
3-Year EPS Expectations	0.241	4.052	0.736	0.183
	(0.217, 0.263)	(2.371, 6.040)	(0.341, 1.176)	(0.100, 0.287)
4-Year EPS Expectations	0.226	4.143	1.537	0.373
	(0.193, 0.251)	(2.406, 6.008)	(0.748, 2.041)	(0.159, 0.667)

This table displays results for the following two-stage least squares regression:

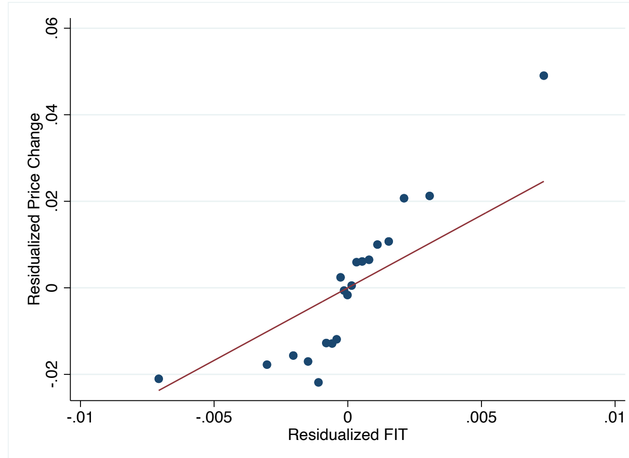
$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$) or quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). FE_t are quarter fixed effects. All units are in percentage points (i.e. 1.0 is 1%). Point estimates are the medians of quarterly block-bootstrapped sampling distributions with 100 samples. 95% confidence intervals from the block-bootstrapped sampling distributions are reported in parentheses. The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

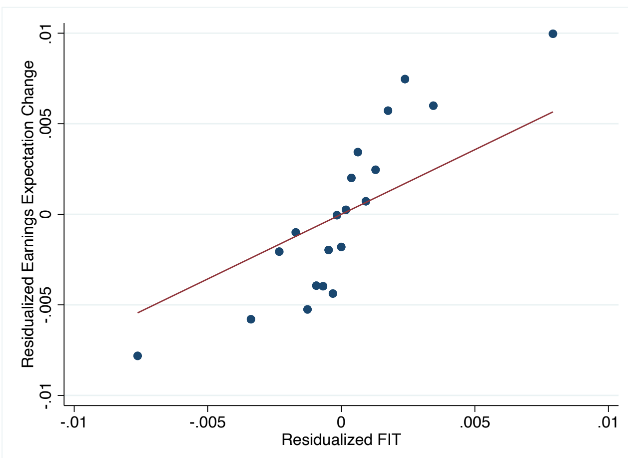
Figure D2: Binscatter Plots for First Stage and Reduced Form



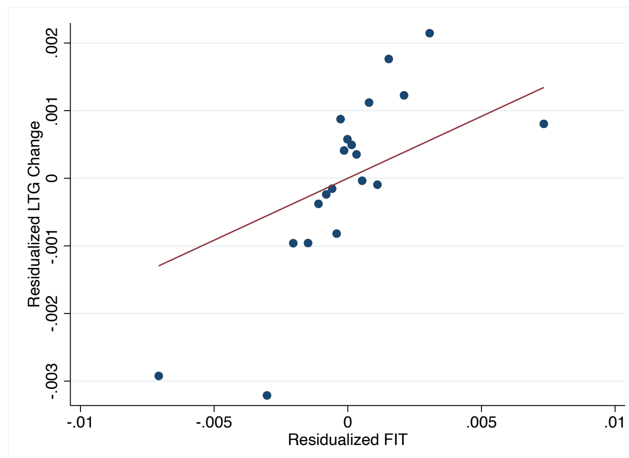
(a) First-Stage (Annual EPS Expectation Revisions)



(b) First-Stage (LTG)



(c) Reduced-Form (Annual EPS Expectation Revisions)



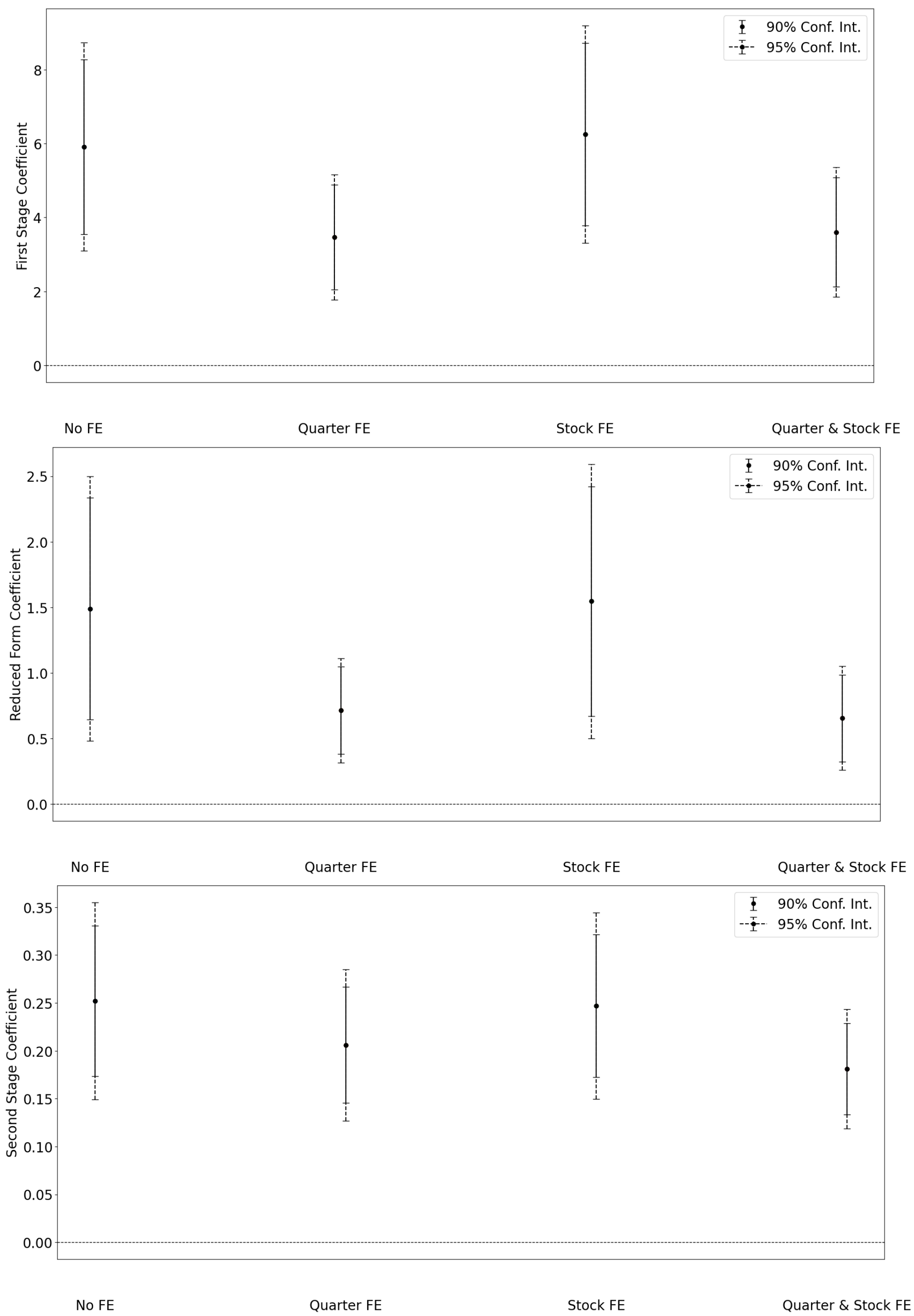
(d) Reduced-Form (LTG)

This figure displays binscatter plots for the following first-stage and reduced-form regressions:

$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + b_1 \text{FIT}_{n,t} + FE_t + e_{2,a,n,t},$$

The first stage (Panels (a) and (b)) regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$) for both the quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) sample in Panel (a) and the changes in LTG expectations ($\Delta \text{LTG}_{a,n,t}$) sample in Panel (b). The reduced form (Panels (c) and (d)) regresses quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) in Panel (c) and changes in LTG expectations ($\Delta \text{LTG}_{a,n,t}$) in Panel (d) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). FE_t are quarter fixed effects. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Figure D3: Alternative FIT Specifications — Annual EPS Expectations

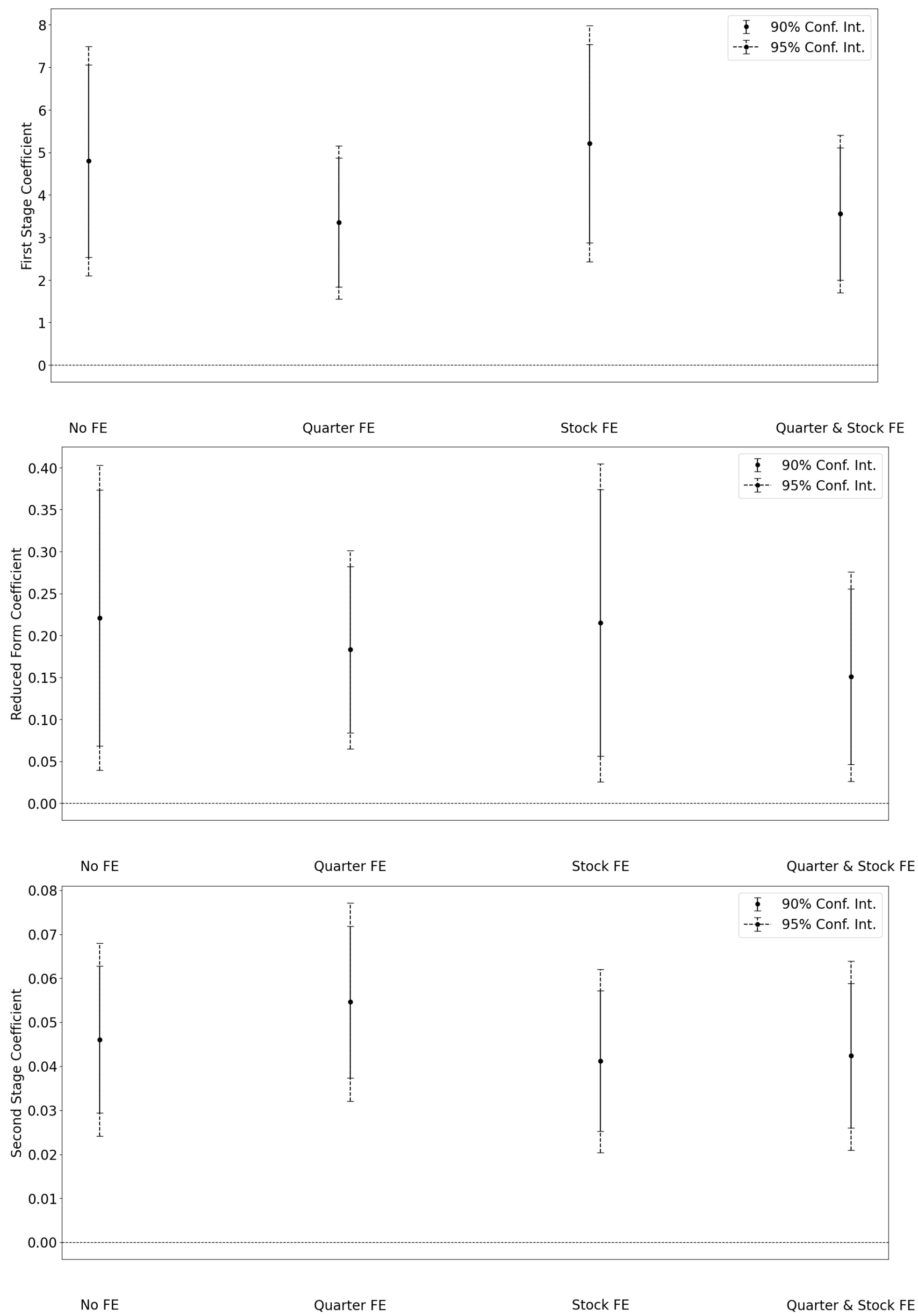


This figure displays results for the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + \boldsymbol{X}_{n,t} + e_{1,a,n,t}$$
$$\Delta E_{a,n,t+4h|t} = b_0 + b_1 \mathrm{FIT}_{n,t} + \boldsymbol{X}_{n,t} + e_{2,a,n,t,h}.$$

The first stage regresses quarterly percent price changes ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). The reduced form regresses quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). $\boldsymbol{X}_{n,t}$ potentially includes stock and quarter fixed effects. The solid error bars display 90% confidence intervals, while the dashed error bars display 95% confidence intervals. The 90% and 95% confidence intervals for the second-stage coefficients using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1982-04:2020-12.

Figure D4: Alternative FIT Specifications — LTG Expectations

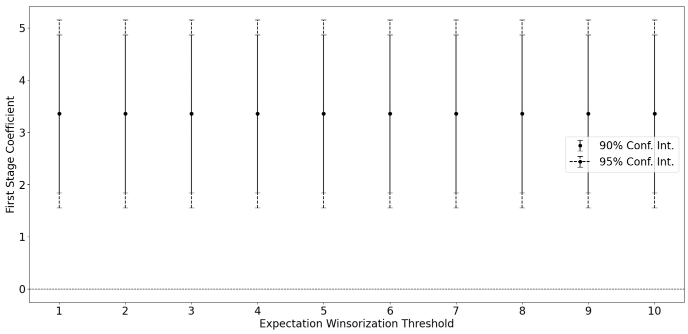


This figure displays results for the following two-stage least squares regression:

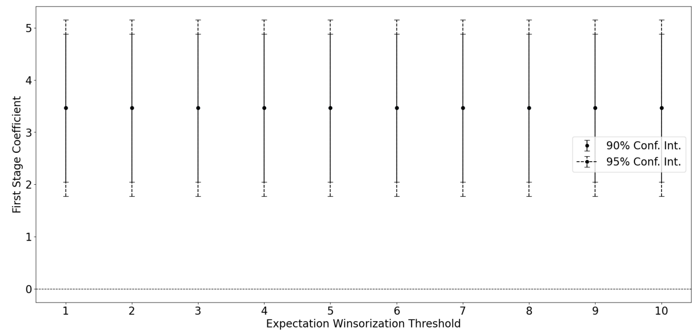
$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + \boldsymbol{X}_{n,t} + e_{1,a,n,t}$$
$$\Delta \mathrm{LTG}_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{X}_{n,t} + e_{2,a,n,t},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). The second stage regresses quarterly changes in LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). $\boldsymbol{X}_{n,t}$ potentially includes stock and quarter fixed effects. The solid error bars display 90% confidence intervals, while the dashed error bars display 95% confidence intervals. The 90% and 95% confidence intervals for the second-stage coefficients using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12.

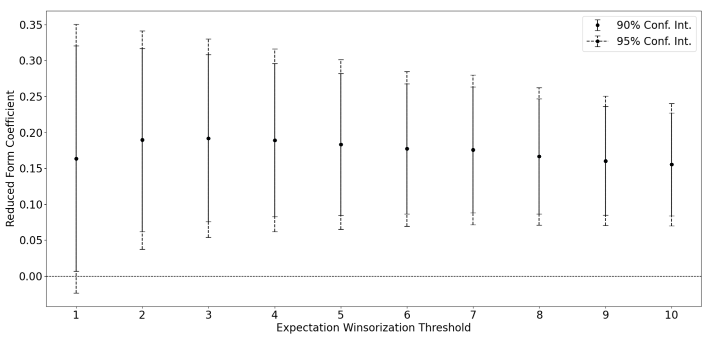
Figure D5: Effect of Prices on Analyst Expectations Using FIT as Instrument — Alternative Winsorizations



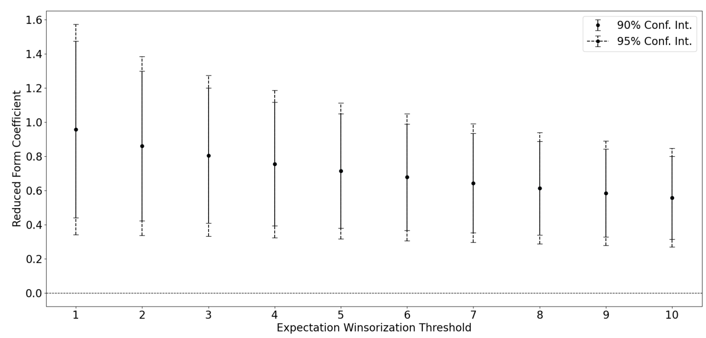
(a) LTG Expectations (First Stage)



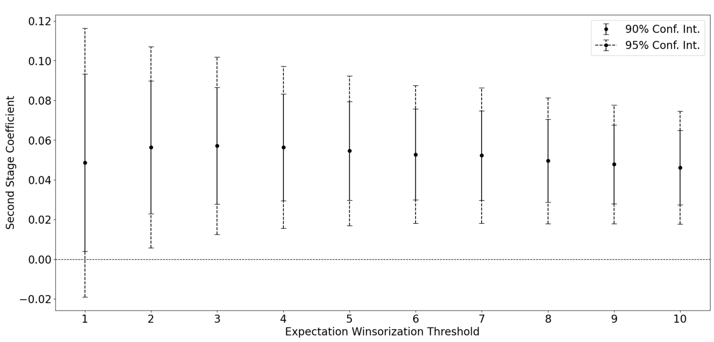
(b) Annual EPS Expectations (First Stage)



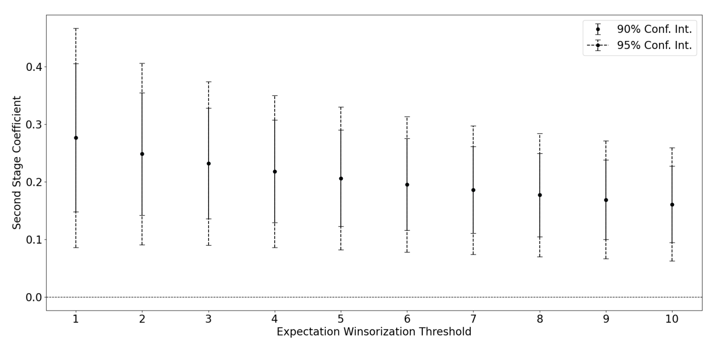
(c) LTG Expectations (Reduced Form)



(d) Annual EPS Expectations (Reduced Form)



(e) LTG Expectations (2SLS)



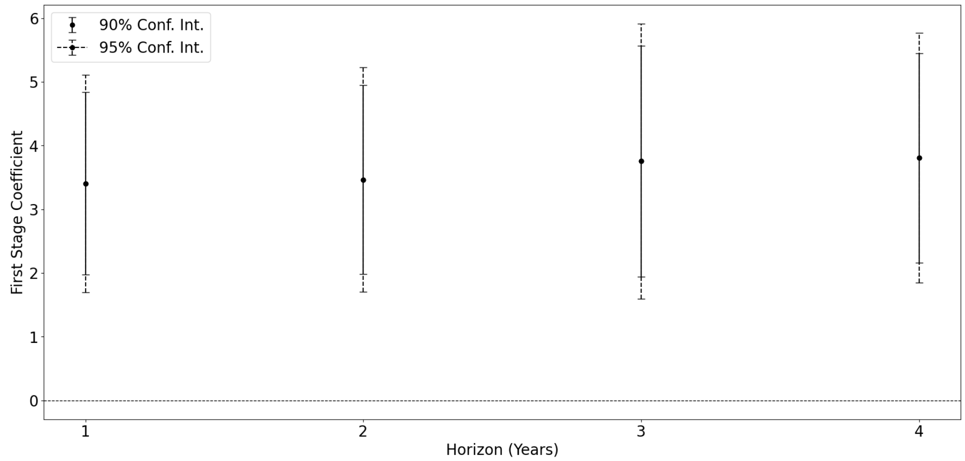
(f) Annual EPS Expectations (2SLS)

This table displays results for the following two-stage least squares regression:

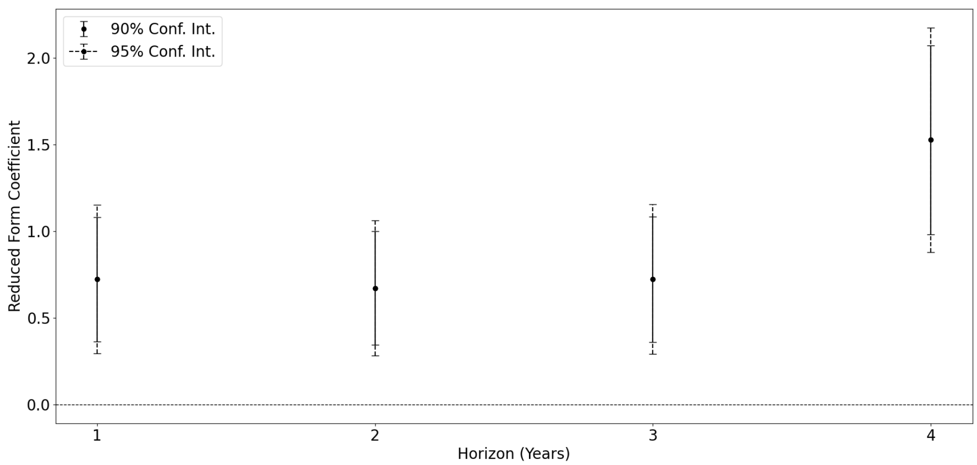
$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \text{LTG}_{a,n,t}$) or quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). FE_t are quarter fixed effects. Panels (a), (c), and (e) display results for the LTG expectations sample. Panels (b), (d), and (f) display results for the annual EPS expectations sample. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. Panels (e) and (f) report confidence intervals using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectations.

Figure D6: First-Stage and Reduced-Form Regressions for Annual EPS Expectations Revisions



(a) First-Stage Results



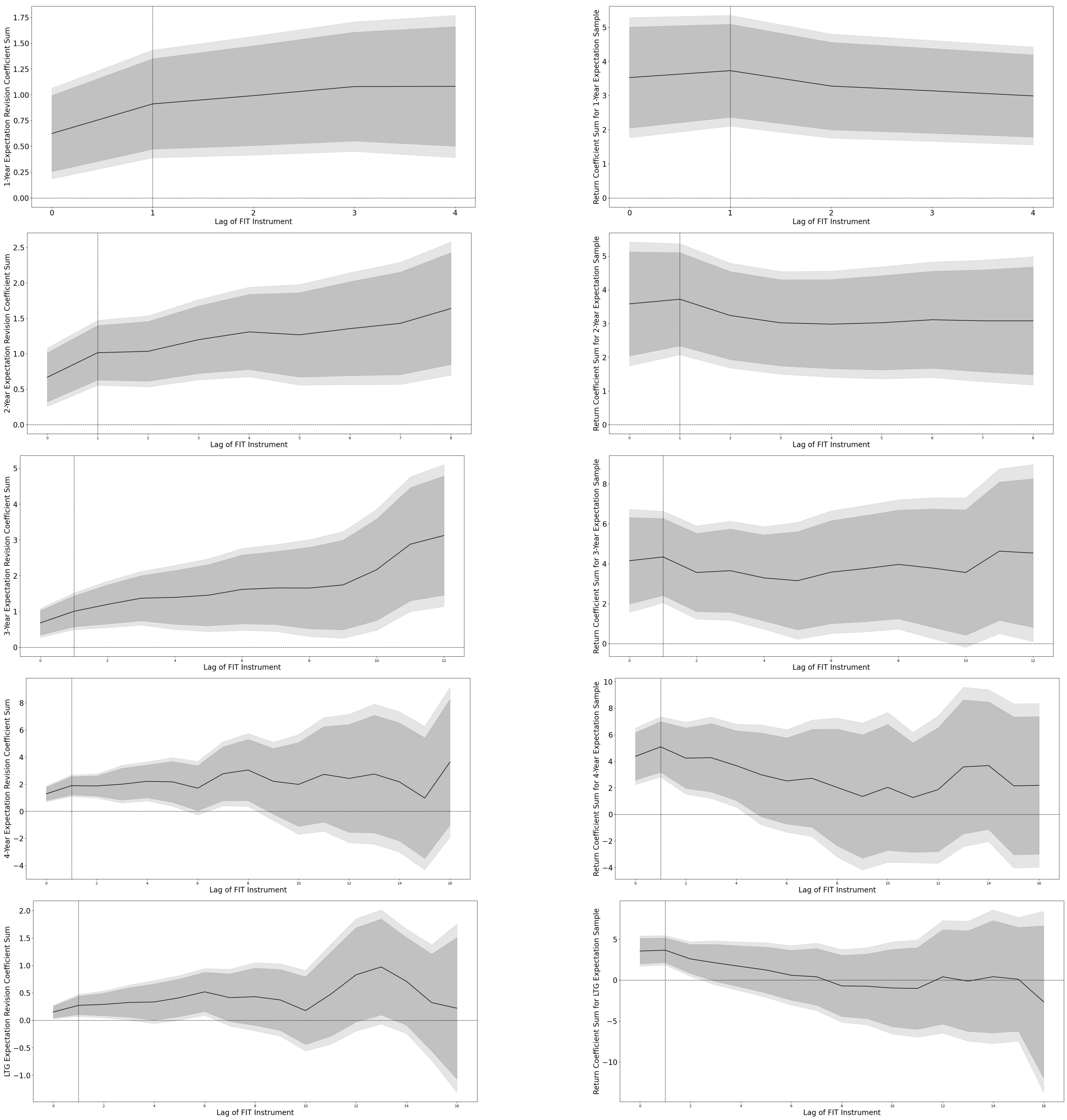
(b) Reduced-Form Results

This figure displays the following first-stage and reduced-form regression results:

$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta E_{a,n,t+4h|t} = b_0 + b_1 \text{FIT}_{n,t} + FE_t + e_{2,a,n,t}.$$

The first stage (in Panel (a)) regresses quarterly percent price changes ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). The reduced form (in Panel (b)) regresses quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). I run this regression for each horizon h from 1 to 4 years. FE_t are quarter fixed effects. Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. The time period is 1982-04:2020-12.

Figure D7: Reversal in Impact of FIT on Analyst Cash Flow Expectations at Longer Horizons

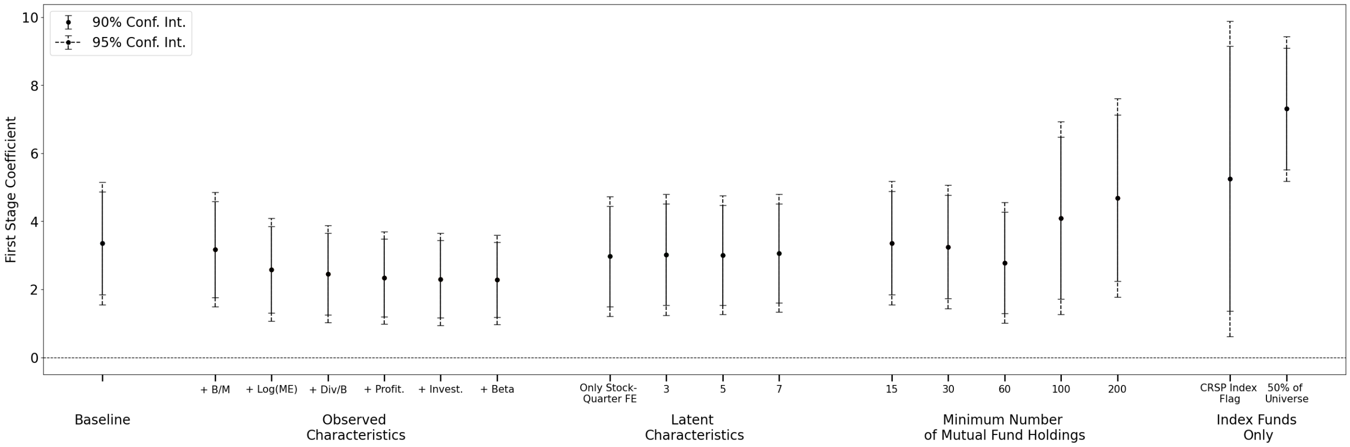


This figure displays the coefficient sums $\sum_{s=0}^{L} \beta_s, L = 1, \ldots, 4$, from the following regression:

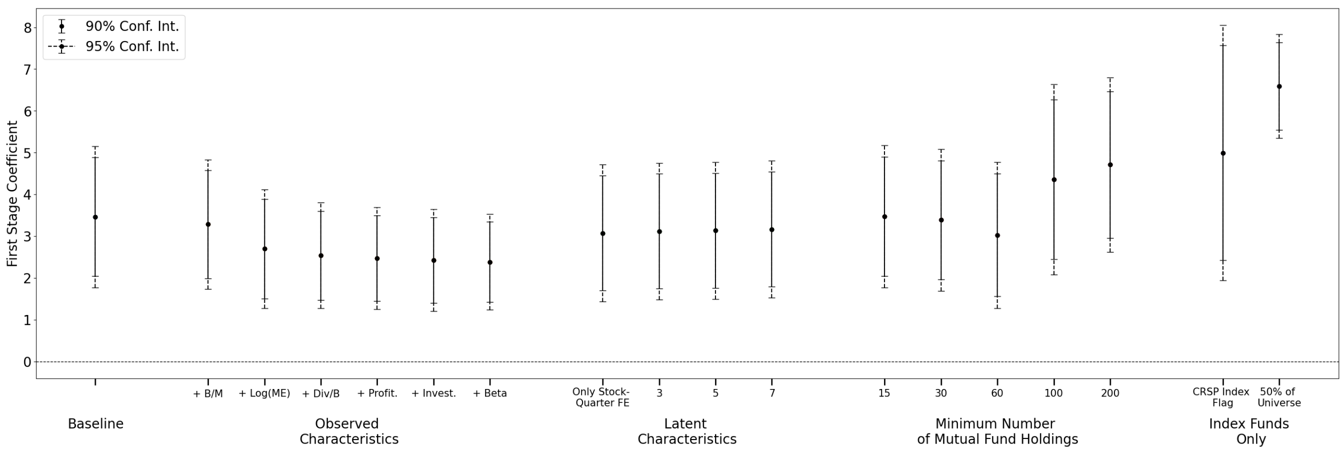
$$\Delta y_{a,n,t} = \sum_{s=0}^{L} \beta_s \text{FIT}_{n,t-s} + FE_t + FE_n + e_{a,n,t}.$$

$\Delta y_{a,n,t}$ is either the quarter-over-quarter change in LTG expectations for analyst institution a for stock n in quarter t $\Delta\text{LTG}_{a,n,t}$, or the quarterly revisions to annual EPS expectations with horizons of one to four years $\Delta E_{a,n,t+4h|t}$. FE_t and FE_n are quarter and stock fixed effects. For EPS expectation horizon h years, I use a maximum lag of $4h$ quarters. For the LTG expectations I use a maximum lag of 16 quarters. Dark and light shaded areas represent 90% and 95% confidence intervals, respectively. Standard errors are clustered by quarter and stock. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Figure D8: Robustness Checks for Using FIT Instrument — First Stage



(a) LTG Expectations



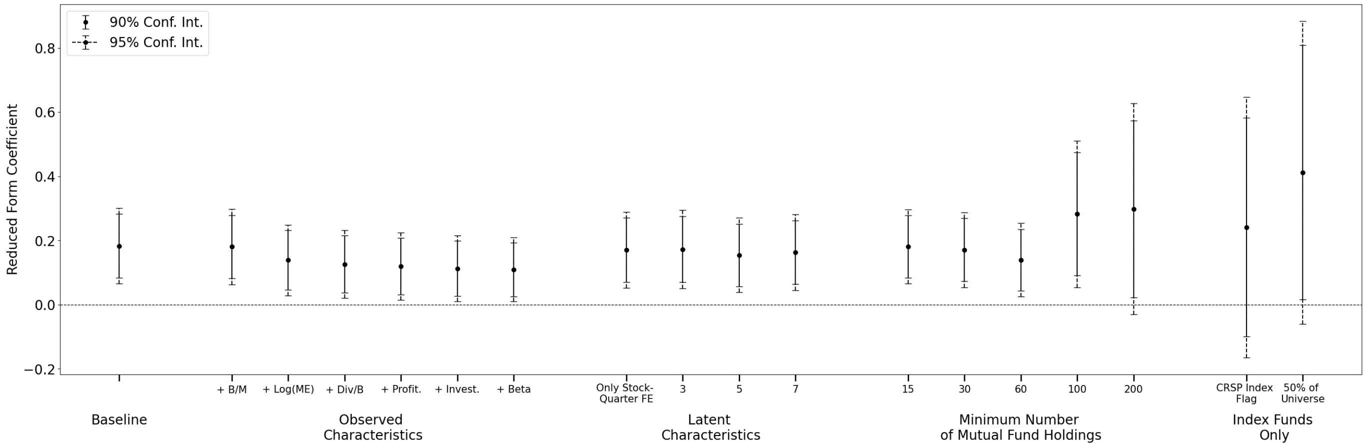
(b) Annual EPS Expectations

This figure displays the following first-stage regression:

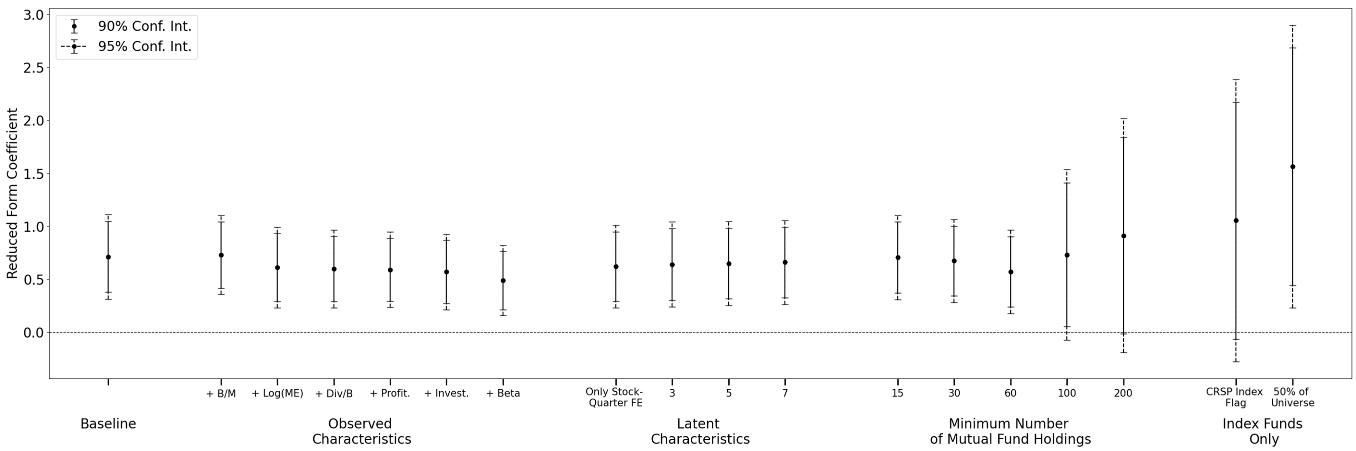
$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t},$$

which regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$) for the LTG expectations sample ($\Delta \text{LTG}_{a,n,t}$ in Panel (a)) and the quarterly revisions to annual EPS expectations with horizons of one to four years sample ($\Delta E_{a,n,t+4h|t}$ in Panel (b)). $\boldsymbol{X}_{n,t}$ includes quarter indicators interacted with either observed (in the "Observed Characteristics" specifications) or latent (in the "Latent Characteristics" specifications) stock characteristics. For the "Observed Characteristics" specifications, each subsequent column adds a control variable (e.g. the right-most column represents the results of the regression with all six control variables). The time period for the "Observed Characteristics" and "Latent Characteristics" specifications is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions. For the "Minimum Number of Mutual Fund Holdings" specifications, each column constructs $\text{FIT}_{n,t}$ only from mutual funds that have at least M holdings, where M is labeled on the x-axis. The time period for the "Minimum Number of Mutual Fund Holdings" specifications is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions. The "Index Funds Only" specifications construct $\text{FIT}_{n,t}$ from only mutual funds identified as index funds by one of the two criteria listed on the x-axis. The time period for the "Index Funds Only" specifications is 1984-09:2020-12. FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%).

Figure D9: Robustness Checks for Using FIT Instrument — Reduced Form



(a) LTG Expectations



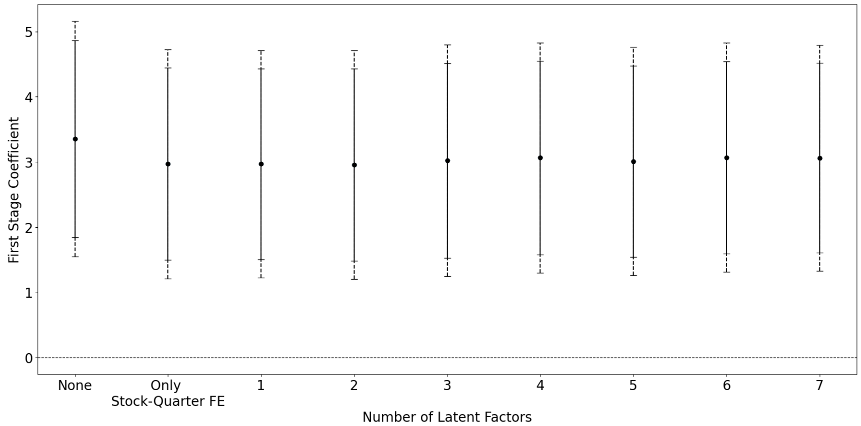
(b) Annual EPS Expectations

This figure displays the following reduced-form regression:

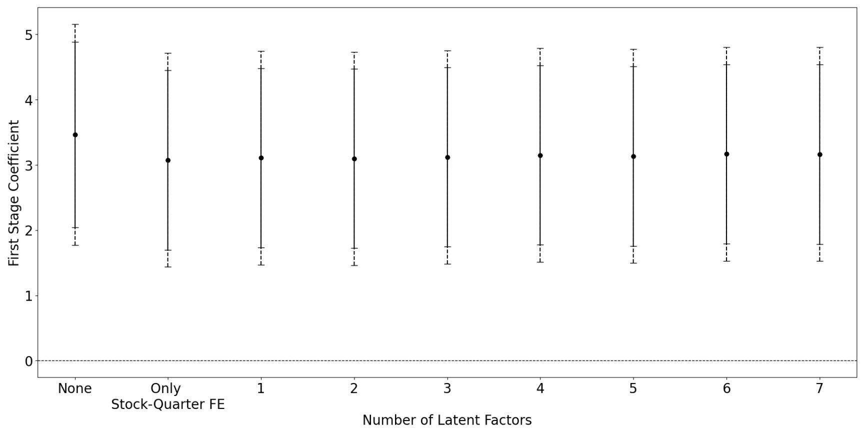
$$\Delta y_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t},$$

which regresses changes in LTG expectations ($\Delta\text{LTG}_{a,n,t}$ in Panel (a)) and quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$ in Panel (b)) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). $\boldsymbol{X}_{n,t}$ includes quarter indicators interacted with either observed (in the "Observed Characteristics" specifications) or latent (in the "Latent Characteristics" specifications) stock characteristics. For the "Observed Characteristics" specifications, each subsequent column adds a control variable (e.g. the right-most column represents the results of the regression with all six control variables). The time period for the "Observed Characteristics" and "Latent Characteristics" specifications is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions. For the "Minimum Number of Mutual Fund Holdings" specifications, each column constructs $\text{FIT}_{n,t}$ only from mutual funds that have at least M holdings, where M is labeled on the x-axis. The time period for the "Minimum Number of Mutual Fund Holdings" specifications is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions. The "Index Funds Only" specifications construct $\text{FIT}_{n,t}$ from only mutual funds identified as index funds by one of the two criteria listed on the x-axis. The time period for the "Index Funds Only" specifications is 1984-09:2020-12. FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%).

Figure D10: FIT First Stage Regressions — Latent Characteristics Controls



(a) Changes in LTG Expectations



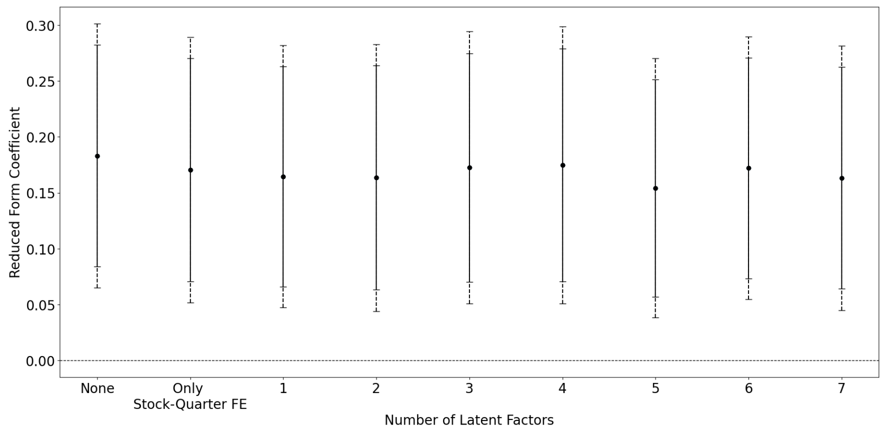
(b) Revisions in Annual EPS Expectations

This figure displays the following regression:

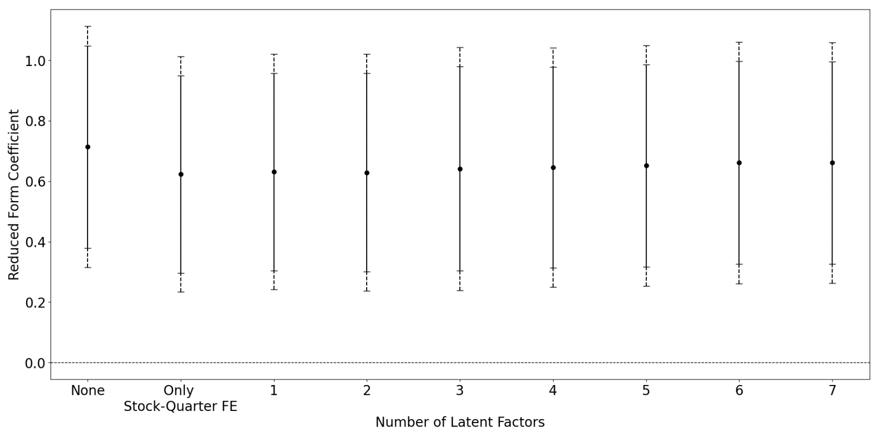
$$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t},$$

which regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\text{FIT}_{n,t}$). Panels (a) and (b) displays the regression results for the LTG expectations ($\Delta\text{LTG}_{a,n,t}$) and annual EPS expectations ($\Delta E_{a,n,t+4h|t}$) samples, respectively. $\boldsymbol{X}_{n,t}$ includes quarter indicators interacted with the latent stock-quarter fixed effect and up to seven latent stock characteristics estimated from the latent factor model described in Section 5.4.2. FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Figure D11: FIT Reduced-Form Regressions — Latent Characteristics Controls



(a) Changes in LTG Expectations



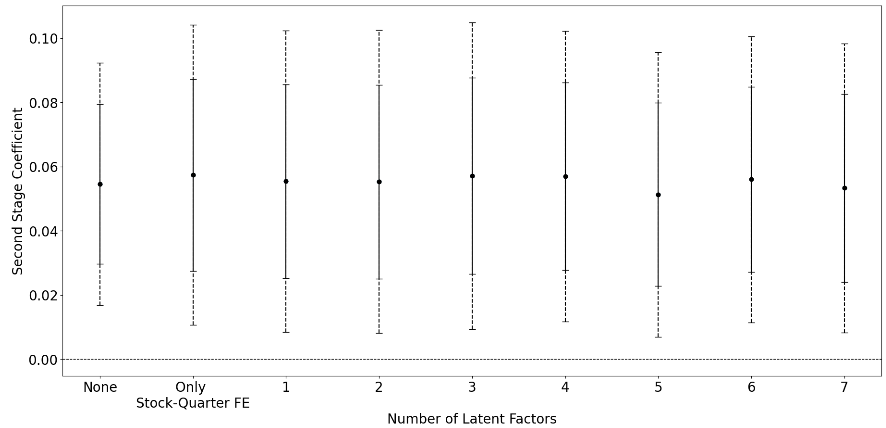
(b) Revisions in Annual EPS Expectations

This figure displays the following regression:

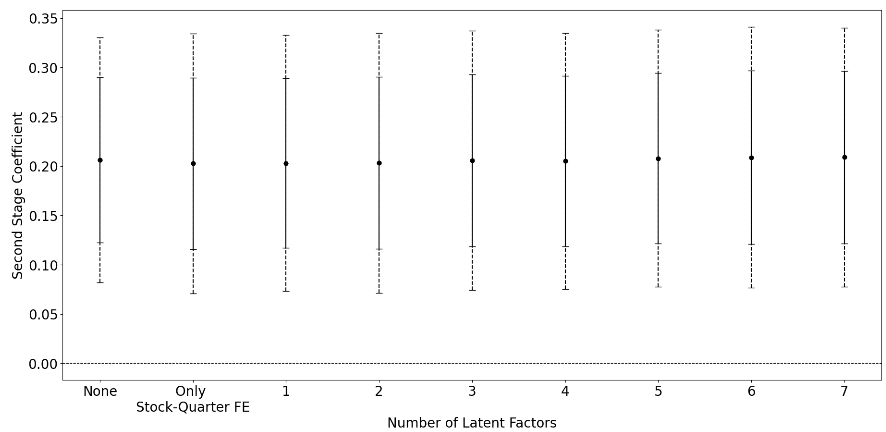
$$\Delta y_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + \boldsymbol{\beta}' \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t},$$

which regresses changes in LTG expectations ($\Delta y_{a,n,t} = \Delta \mathrm{LTG}_{a,n,t}$) in Panel (a) and quarterly revisions in annual EPS expectations ($\Delta y_{a.n,t} = \Delta E_{a,n,t+4h|t}$) in Panel (b) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). $\boldsymbol{X}_{n,t}$ includes quarter indicators interacted with the latent stock-quarter fixed effect and up to seven latent stock characteristics estimated from the latent factor model described in Section 5.4.2. FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Figure D12: FIT Second-Stage Regressions — Latent Characteristics Controls



(a) LTG Expectations



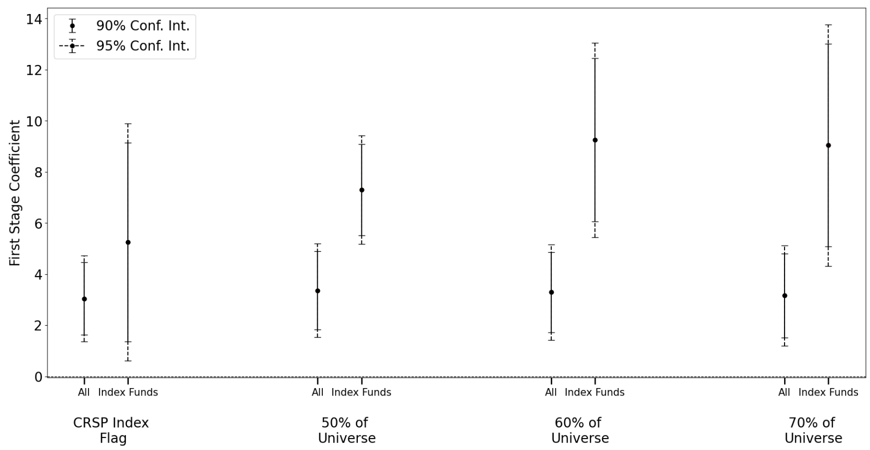
(b) Annual EPS Expectations

This figure displays the following two-stage least squares regression:

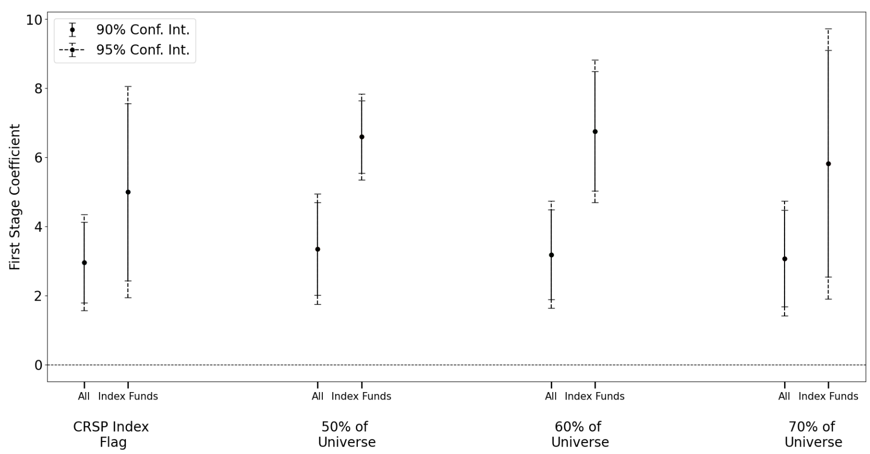
$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + b_1 \mathrm{FIT}_{n,t} + \boldsymbol{\beta}_2^{'} \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t}.$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$) in Panel (a) and quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) in Panel (b) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). $\boldsymbol{X}_{n,t}$ includes quarter indicators interacted with the latent stock-quarter fixed effect and up to seven latent stock characteristics estimated from the latent factor model described in Section 5.4.2. FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals, both of which are computed using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1983-01:2020-12 for the LTG expectations and 1982-04:2020-12 for the annual EPS expectation revisions.

Figure D13: FIT First-Stage Regressions — Passive Funds Only



(a) LTG Expectations



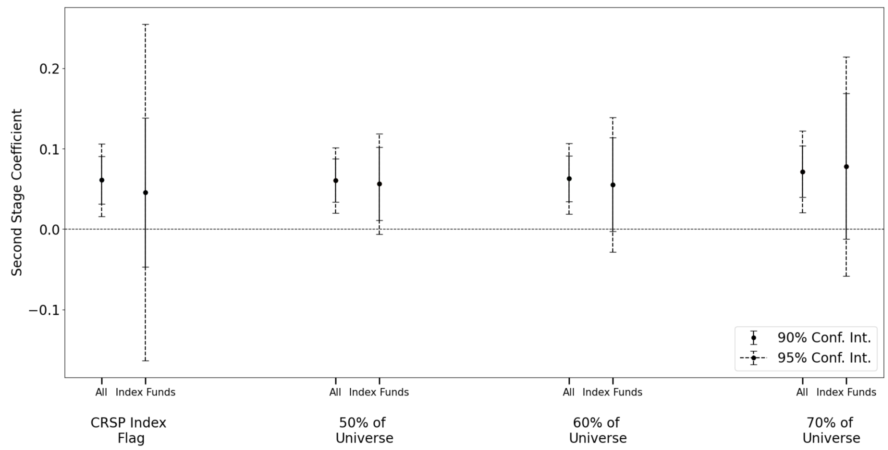
(b) Annual EPS Expectations

This table displays results for the following regression:

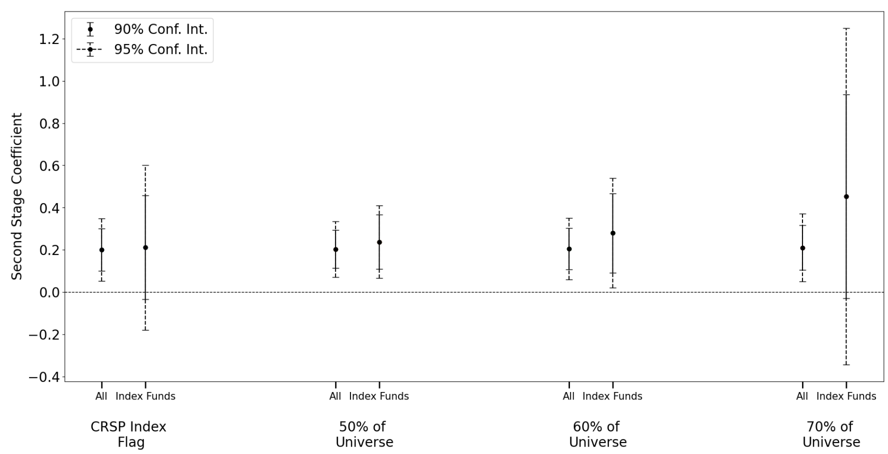
$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + FE_t + e_{1,a,n,t},$$

which regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$) for the LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$) sample in Panel (a) and the quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) sample in Panel (b). FE_t are quarter fixed effects. For each passive fund classification on the x-axis, both the results using the FIT instrument constructed from only passive funds and the results using the standard FIT instrument in the matched sample are displayed (i.e. stock n will be dropped from the sample if $\mathrm{FIT}_{n,t}$ cannot be constructed using only passive funds because no passive funds hold stock n). The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals, both of which are computed using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1984-09:2020-12. Classification accuracy based on the CRSP index fund flag and "target date" name definition of "passive" for the 50% of Universe definition is 84% with a false positive rate of 8%. Accuracy for the 60% of Universe definition is 87% with a false positive rate of 3%. Accuracy for the 70% of Universe definition is 88% with a false positive rate of 1%.

Figure D14: FIT Reduced-Form Regressions — Passive Funds Only



(a) LTG Expectations



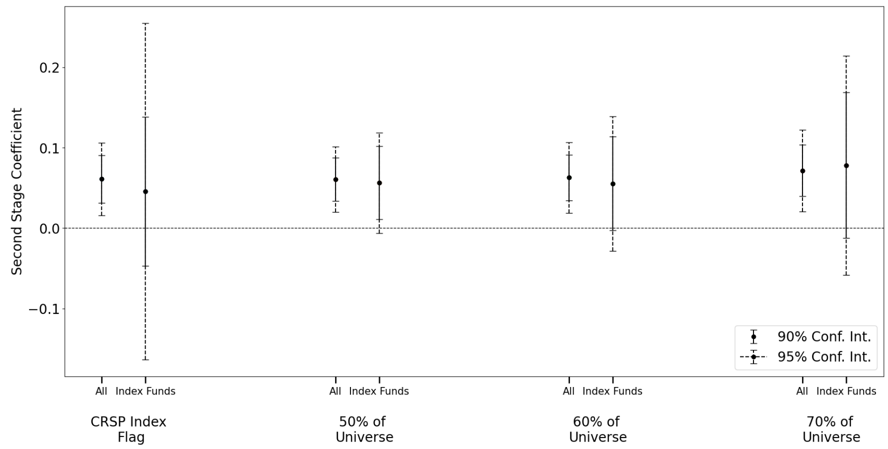
(b) Annual EPS Expectations

This table displays results for the following regression:

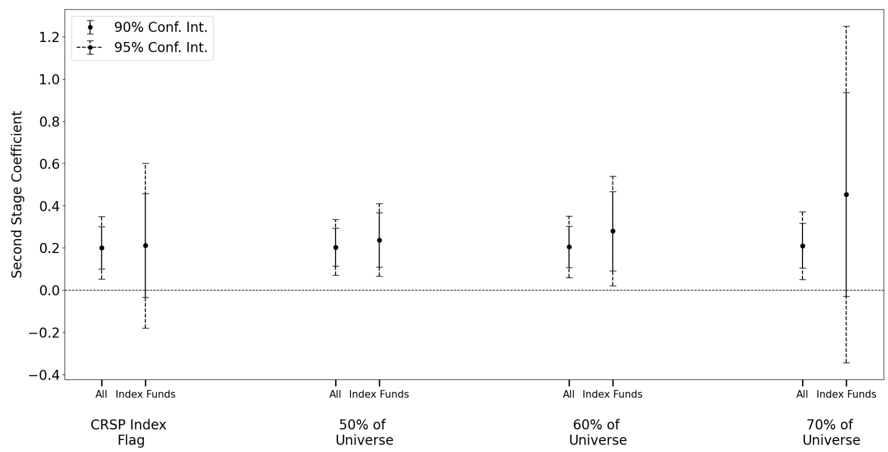
$$\Delta y_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + FE_t + e_{1,a,n,t},$$

of percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$) for the LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$) in Panel (a) and the quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) in Panel (b). FE_t are quarter fixed effects. For each passive fund classification on the x-axis, both the results using the FIT instrument constructed from only passive funds and the results using the standard FIT instrument in the matched sample are displayed (i.e. stock n will be dropped from the sample if $\mathrm{FIT}_{n,t}$ cannot be constructed using only passive funds because no passive funds hold stock n). The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1984-09:2020-12. Classification accuracy based on the CRSP index fund flag and "target date" name definition of "passive" for the 50% of Universe definition is 84% with a false positive rate of 8%. Accuracy for the 60% of Universe definition is 87% with a false positive rate of 3%. Accuracy for the 70% of Universe definition is 88% with a false positive rate of 1%.

Figure D15: FIT Second-Stage Regressions — Passive Funds Only



(a) LTG Expectations



(b) Annual EPS Expectations

This table displays results for the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the flow-induced trading instrument ($\mathrm{FIT}_{n,t}$). The second stage regresses changes in LTG expectations ($\Delta \mathrm{LTG}_{a,n,t}$) in Panel (a) and quarterly revisions to annual EPS expectations with horizons of one to four years ($\Delta E_{a,n,t+4h|t}$) in Panel (b) on the instrumented price changes ($\Delta \hat{p}_{a,n,t}$). For each passive fund classification on the x-axis, both the results using the FIT instrument constructed from only passive funds and the results using the standard FIT instrument in the matched sample are displayed (i.e. stock n will be dropped from the sample if $\mathrm{FIT}_{n,t}$ cannot be constructed using only passive funds because no passive funds hold stock n). FE_t are quarter fixed effects. The solid error bars represent 90% confidence intervals, while the dashed error bars display 95% confidence intervals, both of which are computed using the tF procedure of Lee et al. (2022). Standard errors are clustered by quarter. All units are in percentage points (i.e. 1.0 is 1%). The time period is 1984-09:2020-12. Classification accuracy based on the CRSP index fund flag and "target date" name definition of "passive" for the 50% of Universe definition is 84% with a false positive rate of 8%. Accuracy for the 60% of Universe definition is 87% with a false positive rate of 3%. Accuracy for the 70% of Universe definition is 88% with a false positive rate of 1%.

D.1 Calculating Rotemberg Weights

The FIT instrument is

$$\text{FIT}_{n,t} = \sum_i f_{i,t} S_{i,n,t-2}$$

Let

$$\boldsymbol{S}_{i,t} = [S_{i,n,t}]_{n=1}^{N}$$

be the vector of ownership shares for all stocks for fund i in quarter t. Let $\Delta \boldsymbol{p}_{a,t} = [\Delta p_{a,n,t}]_{n=1}^{N}$, where N is the number of stocks.

Following Goldsmith-Pinkham, Sorkin and Swift (2020), the Rotemberg weight the FIT instrument places on fund i in quarter t is

$$r_{i,t} = \frac{f_{i,t} \sum_a \boldsymbol{S}^{'}_{i,t} \Delta \boldsymbol{p}_{a,t}}{\sum_j f_{j,t} \sum_a \boldsymbol{S}^{'}_{j,t} \Delta \boldsymbol{p}_{a,t}}.$$

I then aggregate the Rotemberg weights up to the fund style level. The aggregate weight for style s is

$$r_s = \sum_{i \in s} \sum_t r_{i,t},$$

where the outer summation sums over all funds i that belong to style s.

Table D6 displays the aggregate Rotemberg weights for all styles.

I use CRSP style codes to assign each fund i in each quarter t to a style. I construct the FIT instrument by merging mutual fund flows from the CRSP Mutual Fund database — in which funds are identified by the CRSP Fund Number identifier — and mutual fund holdings from the Thomson Reuters S12 database — in which funds are identified by Wharton Financial Institution Center Number (WFICN) identifier. I use the WRDS MFLinks database to link these identifiers. Some WFICNs map to multiple CRSP Fund Numbers, and so map to multiple CRSP style codes. In these cases, I use the most common CRSP style code for each (WFICN, quarter) pair. In the few cases where two CRSP style codes are equally common for a particular (WFICN, quarter) pair, I assign the fund to a new category defined by the pair of styles. Additionally, some (WFICN, quarter) pairs are unable to be matched to CRSP style codes and are assigned the category "None" in Table D6; the total Rotemberg weight of these (WFICN, quarter) pairs is small.

Table D6: Rotemberg Weights for All Important Fund Styles

Style	LTG Sample	Style	Annual EPS Expectation Sample
Domestic Equity Growth	82.74	Domestic Equity Small Cap	38.08
Domestic Equity Small Cap	27.23	Domestic Equity Growth	35.97
Domestic Equity Income	3.98	Domestic Equity Mid Cap	10.20
Domestic Equity Growth & Income	2.37	Domestic Equity Micro Cap	4.08
Domestic Equity Micro Cap	1.69	None	2.75
Domestic Equity Mid Cap	1.40	Domestic Equity Large Cap	2.26
Foreign Equity	1.13	Domestic Equity Growth & Income	2.06
Domestic Equity Total Return	0.18	Foreign Equity	1.58
Foreign Equity Regional China	0.16	Foreign Equity Small Cap	1.29
(Domestic Equity Growth , Foreign Equity)	0.05	Domestic Equity Income	1.13
(Domestic Equity Small Cap , Domestic Equity Growth)	0.04	Foreign Equity Regional Emerging Markets	0.29
Foreign Equity Regional Latin America	0.04	Foreign Equity Hedged	0.18
Foreign Equity Regional Pacific	0.01	Foreign Equity Regional China	0.06
(Domestic Equity Growth , Domestic Equity Total Return)	0.01	Domestic Equity Total Return	0.03
(Domestic Equity Mid Cap , Domestic Equity Small Cap)	0.00	(Domestic Equity Growth , Foreign Equity)	0.03
(Domestic Equity Small Cap , Domestic Equity Growth & Income)	0.00	Foreign Equity Regional Latin America	0.02
Foreign Equity Regional India	0.00	(Domestic Equity Small Cap , Domestic Equity Growth)	0.00
(Domestic Equity Growth , Domestic Equity)	-0.00	(Domestic Equity Growth & Income , Domestic Equity Growth)	0.00
Foreign Equity Regional Japan	-0.00	Foreign Equity Regional India	0.00
(Domestic Equity Large Cap , Domestic Equity Growth)	-0.00	(Domestic Equity Mid Cap , Domestic Equity Small Cap)	0.00
(Domestic Equity Mid Cap , Domestic Equity Growth)	-0.00	(Domestic Equity Small Cap , Domestic Equity Growth & Income)	0.00
Foreign Equity Regional Europe	-0.00	Domestic Equity	0.00
Domestic Equity	-0.00	(Domestic Equity Growth , Domestic Equity Total Return)	0.00
Domestic Equity	-0.00	Foreign Equity Regional Japan	0.00
(Domestic Equity Growth & Income , Domestic Equity Growth)	-0.02	Foreign Equity Regional Pacific	0.00
Foreign Equity Regional Pacific ex Japan	-0.02	(Domestic Equity Growth , Domestic Equity)	-0.00
Foreign Equity Growth	-0.08	(Domestic Equity Large Cap , Domestic Equity Growth)	-0.00
Foreign Equity Hedged	-0.23	Foreign Equity Regional Europe	-0.00
Domestic Equity Large Cap	-0.41	(Domestic Equity Mid Cap , Domestic Equity Growth)	-0.00
Foreign Equity Regional Emerging Markets	-1.03	Foreign Equity Regional Pacific ex Japan	-0.00
Foreign Equity Small Cap	-4.54	Domestic Equity	-0.00
None	-14.70	Foreign Equity Growth	-0.03

Rotemberg weights for all fund styles for both the LTG expectations ($\Delta LTG_{a,n,t}$) and annual EPS expectations ($\Delta E_{a,n,t+4h|t}$) samples.

D.2 Within Fund-Quarter Cross-Sectional Ownership Share Variation Explained

To calculate the percentage of within fund-quarter cross-sectional variance in ownership shares explained by a certain set of stock characteristics $\boldsymbol{X}_{n,t}$, I calculate the within R^2 of the following panel regression:

$$S_{i,n,t} = \boldsymbol{c}_{i,t}^{'}\boldsymbol{X}_{n,t} + FE_{i,t} + \tilde{S}_{i,n,t}$$

$$\text{Within } R^2 = 1 - \frac{\sum_{i,n,t} \tilde{S}_{i,n,t}^2}{\sum_{i,n,t} \left(S_{i,n,t} - \bar{S}_{i,t}\right)}$$

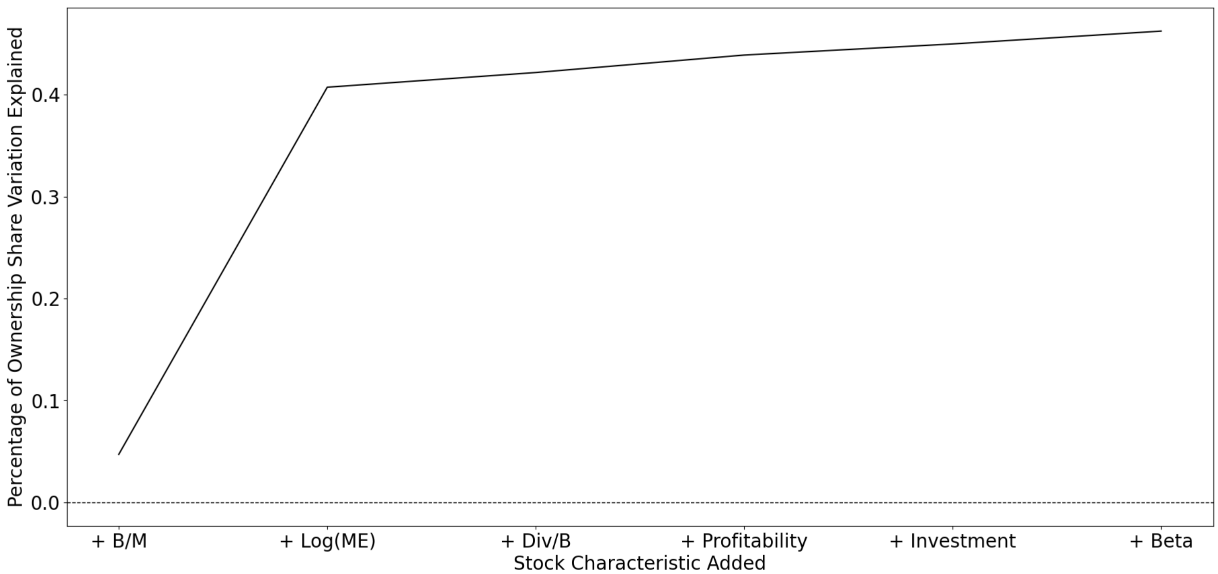
$$\bar{S}_{i,t} = \frac{1}{\text{Num Holdings}_{i,t}} \sum S_{i,n,t},$$

where $\bar{S}_{i,t}$ is the average ownership share for fund i in quarter t. To be clear, the within R^2 does not include the variation in ownership shares explained by the fund-quarter fixed effects.

Figure D16 displays the cumulative percentage of ownership share variance explained by the six observed stock characteristics associated with mutual fund styles: book-to-market ratio, log market equity, dividend-to-book equity ratio, profitability, investment, and market beta. These six characteristics explain a total of 46% of the within fund-quarter cross-sectional variance in ownership shares. Much of that variation is explained by size (log market equity).

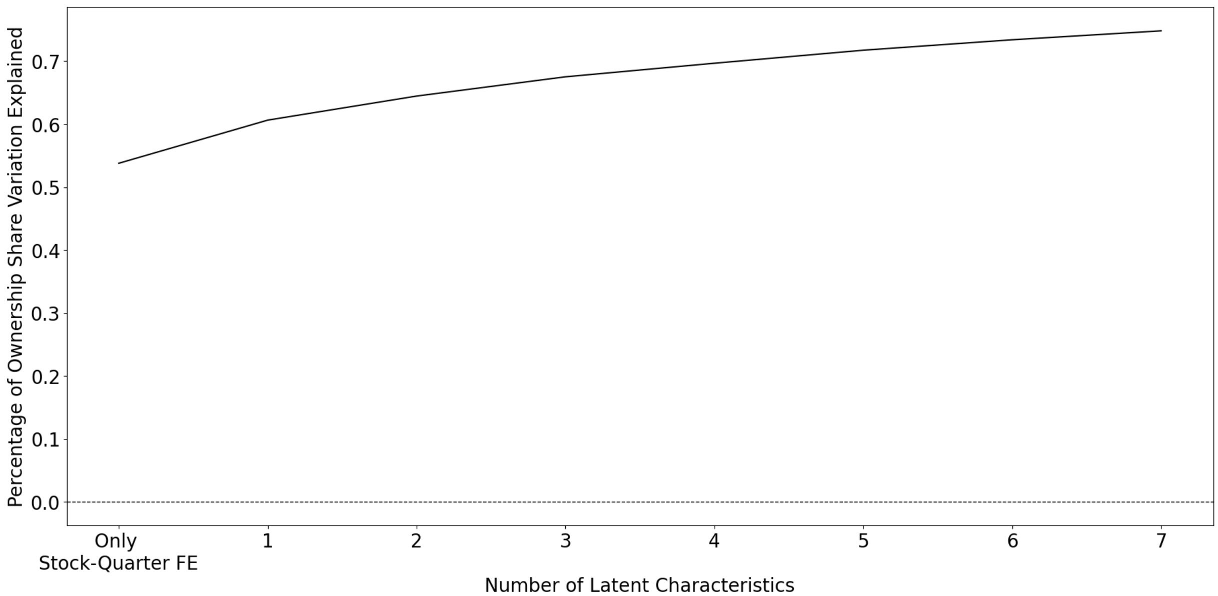
Figure D17 displays the cumulative percentage of ownership share variance explained by the latent stock-quarter fixed effects and seven latent stock characteristics from the latent factor model (23) from Section 5.4.2. These eight latent characteristics explain a total of 75% of the within fund-quarter cross-sectional variance in ownership shares. Much of that variation is explained by the latent stock-quarter fixed effects.

Figure D16: Percentage of Within Fund-Quarter Cross-Sectional Variation in Ownership Shares Explained by Observed Stock Characteristics



Percentage of within fund-quarter cross-sectional variation in ownership shares explained by the six observed stock characteristics associated with mutual funds styles: book-to-market ratio, log market equity, dividend-to-book equity ratio, profitability, investment, and market beta. The graph is cumulative: each column adds an additional stock characteristic to the regression. So the sixth column reports the within R^2 from the regression featuring all six stock characteristics.

Figure D17: Percentage of Within Fund-Quarter Cross-Sectional Variation in Ownership Shares Explained by Latent Stock Characteristics



Percentage of within fund-quarter cross-sectional variation in ownership shares explained by the eight latent stock characteristics: the latent stock-quarter fixed effects and the seven latent stock characteristics from the latent factor model (23) from Section 5.4.2. The graph is cumulative: each column adds an additional latent characteristic to the regression. So the eighth column reports the within R^2 from the regression featuring the latent stock-quarter fixed effects and all seven latent stock characteristics.

D.3 Latent Factor Model in Ownership Shares

I fit latent factor model (23) quarter-by-quarter using the regularized singular value decomposition technique of Funk (2006). This method decomposes the mutual fund-by-stock matrix of ownership shares ($\mathbf{S}_t = [S_{i,n,t}]_{i,n}$) into the product of a matrix of fund-specific factor loadings ($\mathbf{C}_t = \left[\boldsymbol{c}^{'}_{i,t}\right]_i$) with a matrix of stock characteristics ($\mathbf{X}_t = [\boldsymbol{X}_{n,t}]_n$), after removing fund-quarter ($\mathbf{FE}_t^{Fund}$) and stock-quarter ($\mathbf{FE}_t^{Stock}$) fixed effects:

$$\mathbf{S}_t = \mathbf{C}_t\mathbf{X}_t + \mathbf{FE}_t^{Fund} + \mathbf{FE}_t^{Stock} + \tilde{\mathbf{S}}_t,$$

One can estimate matrices $\mathbf{C}_t$, $\mathbf{X}_t$, $\mathbf{FE}_t^{Fund}$, and $\mathbf{FE}_t^{Stock}$ as the minimizers of the following loss function

$$\min_{\mathbf{C}_t,\mathbf{X}_t,\mathbf{FE}_t^{Fund},\mathbf{FE}_t^{Stock}} \sum_{i,n} \left(S_{i,n,t} - \hat{S}_{i,n,t}\right)^2$$
$$\text{s.t. } \hat{S}_{i,n,t} = \boldsymbol{c}^{'}_{i,t}\boldsymbol{X}_{n,t} + FE_{i,t} + FE_{n,t}$$

Empirically, most funds do not hold most stocks. For this reason, I can attain more efficient estimates of $\mathbf{C}_t$, $\mathbf{X}_t$, $\mathbf{FE}_t^{Fund}$, and $\mathbf{FE}_t^{Stock}$ by adding L2 penalties to the least-squares loss function (Funk (2006); Bai and Ng (2019)):

$$\min_{\mathbf{C}_t,\mathbf{X}_t,\mathbf{FE}_t^{Fund},\mathbf{FE}_t^{Stock}} \sum_{i,n} \left(S_{i,n,t} - \hat{S}_{i,n,t}\right)^2 + \gamma_t \left(FE_{i,t}^2 + FE_{n,t}^2 + \|\boldsymbol{c}_{i,t}\|^2 + \|\boldsymbol{X}_{n,t}\|^2\right)$$
$$\text{s.t. } \hat{S}_{i,n,t} = \boldsymbol{c}_{i,t}^{'}\boldsymbol{X}_{n,t} + FE_{i,t} + FE_{n,t}.$$

Since I fit the factor model quarter by quarter, all regularization parameters can vary over time. I conduct three-fold cross-validation within each quarter to choose the quarterly regularization parameter γ_t.

E Details of Covariance Decomposition

Following (2), the proportion of the covariance of prices with analyst cash flow expectations accounted for by the impact of prices on analyst expectations is:

$$\frac{\alpha_h \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y_{a,n,t}^{(h)}\right)}, \tag{39}$$

where $\Delta y_{a,n,t}^{(h)}$ is the change in analyst cash flow expectations for horizon h and $\Delta p_{a,n,t}$ is the contemporaneous change in price. α_h is the impact of price on analyst cash flow expectations for horizon h. This ratio is simply the two-stage least squares estimate of α_h divided by the OLS coefficient from the regression of changes in cash flow expectations on contemporaneous price changes:

$$\begin{aligned} \Delta y_{a,n,t}^{(h)} &= \alpha_h^{OLS}\Delta p_{a,n,t} + e_{a,n,t} \\ \alpha_h^{OLS} &= \frac{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y_{a,n,t}^{(h)}\right)}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}. \end{aligned} \tag{40}$$

However, there might be "measurement error" in the quarterly price changes $\Delta p_{a,n,t}$ that biases the OLS coefficient toward zero. For example, if analysts update cash flow expectations based only on price changes for a subset of days during the quarter (e.g. the price change over the month before the announcement), then the variance of the quarterly price change $\Delta p_{a,n,t}$ exceeds that of this "true" price change $\Delta p_{a,n,t}^T$:

$\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right] > \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]$.

Thus, to calculate the true proportion of $Cov^{CX}\left(\Delta p_{a,n,t}, \Delta E_{a,n,t+4h|t}\right)$ accounted for by this price impact, one must multiply (39) by $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$. I measure this ratio following the method of Pancost and Schaller (2022). Pancost and Schaller (2022) notes that if OLS regression (40) suffers from omitted variable bias (due to common information or sentiment shocks impacting both analyst expectations directly and prices through investor expectations) and measurement error, then α^{OLS} is a linear function of the true parameter α_h:

$$\alpha^{OLS}_h = \frac{\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}\alpha_h + OVB_h.$$

Pancost and Schaller (2022) demonstrate that in an OLS regression of OLS coefficients α^{OLS}_h on two-stage least squares estimates α_h

$$\alpha^{OLS}_h = a + \theta\alpha_h + e_h, \tag{41}$$

θ is a consistent estimator of $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$.[49]

Pancost and Schaller (2022) note that any estimation error in two-stage least squares estimates α_h will create attenuation bias in θ, and so will result in an underestimation of $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$.

Thus, I draw quarterly block-bootstrapped samples, and estimate α^{OLS}_h, α_h, and θ in each sample. I then compute the corrected proportion of covariance explained for each horizon:

$$\frac{\alpha_h \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y^{(h)}_{a,n,t}\right)}\theta \tag{42}$$

in each sample. I also report this proportion for the pooled set of all one to four year EPS expectations. When running regression (41), I constrain θ from above so that $\theta \leq 1$ and the corrected covariance proportion (42) does not exceed 1 for any horizon. I allow θ to be arbitrarily small.

Table E7 displays the covariance proportion explained by the impact of prices on analyst cash flow expectations calculated using the ΔBMI instrument around Russell index reconstitutions. Column 1 reports the unadjusted covariance proportion (39). Column 2 reports the corrected covariance proportion (42), using the estimate $\theta = \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ from regression (41) with no constraints. Column 3 reports the corrected covariance proportion (42), using θ from regression (41) with the upper bound constraint on

[49]Specifically, Pancost and Schaller (2022) demonstrate θ is a consistent estimator of $\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ in the case of classical measurement error. In this situation, $\Delta p_{a,n,t}$ might feature non-classical measurement error as compared to $\Delta p^T_{a,n,t}$. Nevertheless, as discussed in Appendix E.1, the estimator from (42) will still accurately estimate the corrected covariance decomposition (17) even in this non-classical measurement error case.

θ such that $\theta \leq 1$ and the covariance proportion does not exceed one for any horizon (which is a tighter constraint). There is no lower bound constraint on θ. The upper bound constraint on θ does not impact the point estimates but tightens standard errors a little.

Using the ΔBMI instrument, the proportion of covariance explained for the pooled set of all one to four year EPS expectations is 49%, which is similar to (and well within the 95% confidence interval of) the headline 40% reported in Figure 1 Panel (b) using the FIT instrument. However, the standard errors are larger than in Figure 1 Panel (b) because the horizon-specific α_h estimates using the ΔBMI instrument in Table B3 are much noisier than their counterparts in Table 4 and Figure 5 using the FIT instrument. The proportion of covariance explained for the LTG expectations is 4% and not statistically distinguishable from zero, which is consistent with 1) the non-significant estimate of α for the LTG expectations using the ΔBMI instrument in Table B3, and 2) the fact that noisier estimates of the horizon-specific two-stage least squares α_h lead to underestimation of θ.

Table E8 displays the same covariance proportions calculated using the FIT instrument.

Table E7: Proportion of Covariance Explained by Price Impact on Analyst Cash Flow Expectations — ΔBMI Instrument

	Unadjusted	Adjusted - Unconstrained	Adjusted - Constrained
LTG	0.334	0.040	0.040
	(-2.287, 2.338)	(-0.710, 0.779)	(-0.710, 0.779)
Pooled Annual EPS Expectations	1.561	0.488	0.488
	(0.974, 2.348)	(0.018, 0.824)	(0.018, 0.808)
1-Year EPS Expectations	1.765	0.568	0.568
	(0.975, 2.729)	(0.022, 0.868)	(0.022, 0.851)
2-Year EPS Expectations	1.477	0.435	0.435
	(0.887, 2.094)	(0.017, 0.860)	(0.017, 0.822)
3-Year EPS Expectations	0.506	0.158	0.158
	(-1.302, 1.646)	(-0.113, 0.687)	(-0.113, 0.631)
4-Year EPS Expectations	0.842	0.325	0.325
	(-1.808, 3.976)	(-0.095, 0.871)	(-0.095, 0.860)
θ		0.283	0.283
		(0.012, 0.585)	(0.012, 0.569)

Proportion of covariance of changes in cash flow expectations and contemporaneous price changes explained by price impact on analyst cash flow expectations using ΔBMI instrument around Russell index reconstitutions. Point estimates are medians of quarterly block-bootstrapped sampling distributions (100 samples). 95% percent confidence intervals from block-bootstrapped sampling distributions are displayed in parentheses. Column 1 reports covariance proportion (39), which is unadjusted for measurement error. Column 2 reports covariance proportion (42), which is adjusted for measurement error using the estimate $\theta = \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ from regression (41) with no constraints. Column 3 reports covariance proportion (42), which is adjusted for measurement error using the estimate $\theta = \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right] / \mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ from regression (41) with the upper bound constraint on θ such that $\theta \leq 1$ and the covariance proportion does not exceed one for any horizon (there is no lower bound constraint on θ).

Table E8: Proportion of Covariance Explained by Price Impact on Analyst Cash Flow Expectations — FIT Instrument

	Unadjusted	Adjusted - Unconstrained	Adjusted - Constrained
LTG	1.195	0.610	0.610
	(0.691, 1.656)	(0.148, 1.629)	(0.148, 1.000)
Pooled Annual EPS Expectations	0.760	0.412	0.397
	(0.436, 1.017)	(0.107, 0.989)	(0.108, 0.732)
1-Year EPS Expectations	0.757	0.402	0.386
	(0.428, 1.052)	(0.102, 1.076)	(0.102, 0.784)
2-Year EPS Expectations	0.712	0.388	0.381
	(0.395, 0.974)	(0.096, 0.953)	(0.096, 0.703)
3-Year EPS Expectations	0.769	0.404	0.397
	(0.446, 1.182)	(0.127, 1.095)	(0.127, 0.714)
4-Year EPS Expectations	1.649	0.879	0.801
	(0.719, 2.893)	(0.401, 1.280)	(0.396, 1.000)
θ		0.530	0.531
		(0.155, 1.708)	(0.155, 1.000)

Proportion of covariance of changes in cash flow expectations and contemporaneous price changes explained by price impact on analyst cash flow expectations using FIT instrument. Point estimates are medians of quarterly block-bootstrapped sampling distributions (100 samples). 95% percent confidence intervals from block-bootstrapped sampling distributions are displayed in parentheses. Column 1 reports covariance proportion (39), which is unadjusted for measurement error. Column 2 reports covariance proportion (42), which is adjusted for measurement error using the estimate $\theta = \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]/\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ from regression (41) with no constraints. Column 3 reports covariance proportion (42), which is adjusted for measurement error using the estimate $\theta = \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]/\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]$ from regression (41) with the upper bound constraint on θ such that $\theta \leq 1$ and the covariance proportion does not exceed one for any horizon (there is no lower bound constraint on θ).

E.1 Non-Classical Measurement Error Case

It does not matter if the "measurement error" in the observed quarterly price changes is classical measurement error or not. The estimator from (42) will still accurately estimate the corrected covariance decomposition (17) in either case.

Assume the true model is:

$$\Delta p^T_{a,n,t} = M z_{a,n,t} + \epsilon_{a,n,t}$$

$$\Delta y_{a,n,t} = \alpha \Delta p^T_{a,n,t} + \nu_{a,n,t},$$

where $\Delta p^T_{a,n,t}$ is the true price change analysts respond to that I as the outside econometrician do not observe. Instead I observe the full quarterly price change

$$\Delta p_{a,n,t} = \Delta p^T_{a,n,t} + \Delta p^F_{a,n,t},$$

which is the true price change $\Delta p^T_{a,n,t}$ plus the price changes from other days in the quarter that analysts do not learn from $\Delta p^F_{a,n,t}$. I assume $\Delta p^T_{a,n,t}$ and $\Delta p^F_{a,n,t}$ are uncorrelated because there is little within-quarter serial correlation in returns:

$$\mathbb{E}\left[\Delta p^T_{a,n,t} \Delta p^F_{a,n,t}\right] = 0.$$

$\Delta p^F_{a,n,t}$ creates measurement error. However, this measurement error may be correlated with the instrument $z_{a,n,t}$, which may affect all daily price changes in the quarter, not only those daily price changes that analysts respond to:

$$\mathbb{E}\left[z_{a,n,t} \Delta p^F_{a,n,t}\right] \neq 0.$$

Both prices and analyst expectations experience other, possibly correlated shocks ($\epsilon_{a,n,t}$ and $\nu_{a,n,t}$). For example, public, non-price signals about cash flows (e.g. earnings announcements) that both investors and analysts learn from would appear in $\epsilon_{a,n,t}$ and $\nu_{a,n,t}$.

The instrument is uncorrelated with other determinants of analyst expectations $\nu_{a,n,t}$:

$$\mathbb{E}\left[z_{a,n,t} \nu_{a,n,t}\right] = 0.$$

The OLS regression coefficient from

$$\Delta y_{a,n,t} = \alpha^{OLS} \Delta p_{a,n,t} + e_{1,a,n,t},$$

is (assume all variables have been demeaned for simplicity):

$$\begin{aligned}
\alpha^{OLS} &= \frac{\mathbb{E}\left[\Delta y_{a,n,t} \Delta p_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]} \\
&= \frac{\mathbb{E}\left[\left(\alpha \Delta p^T_{a,n,t} + \nu_{a,n,t}\right)\left(\Delta p^T_{a,n,t} + \Delta p^F_{a,n,t}\right)\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]} \\
&= \underbrace{\alpha \frac{\mathbb{V}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]}}_{\text{Attenuation}} + \underbrace{\frac{\mathbb{E}\left[\Delta p^T_{a,n,t} \nu_{a,n,t}\right] + \mathbb{E}\left[\Delta p^F_{a,n,t} \nu_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]}}_{\text{Omitted Variable Bias}}.
\end{aligned}$$

This OLS coefficient α^{OLS} suffers from attenuation due to measurement error ($\Delta p_{a,n,t}$ is not $\Delta p^T_{a,n,t}$) and omitted variable bias (due to the correlation of $\Delta p_{a,n,t}$ with other determinants of analyst expectations $\nu_{a,n,t}$).

The two-stage least squares regression

$$\Delta p_{a,n,t} = a z_{a,n,t} + e_{2,a,n,t}$$
$$\Delta y_{a,n,t} = \alpha^{2SLS} \Delta \hat{p}_{a,n,t} + e_{3,a,n,t},$$

estimates

$$\begin{aligned}
\alpha^{2SLS} &= \frac{\mathbb{E}\left[z_{a,n,t}\Delta y_{a,n,t}\right]}{\mathbb{E}\left[z_{a,n,t}\Delta p_{a,n,t}\right]} \\
&= \frac{\mathbb{E}\left[z_{a,n,t}\left(\alpha\Delta p^T_{a,n,t} + \nu_{a,n,t}\right)\right]}{\mathbb{E}\left[z_{a,n,t}\left(\Delta p^T_{a,n,t} + \Delta p^F_{a,n,t}\right)\right]} \\
&= \alpha\frac{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]}{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right] + \mathbb{E}\left[z_{a,n,t}\Delta p^F_{a,n,t}\right]}.
\end{aligned}$$

This 2SLS coefficient α^{2SLS} suffers from attenuation (i.e. is biased downward[50]) due to measurement error ($\Delta p_{a,n,t}$ is not $\Delta p^T_{a,n,t}$). In the classical measurement error case, $\mathbb{E}\left[z_{a,n,t}\Delta p^F_{a,n,t}\right] = 0$ and so α^{2SLS} does *not* suffer from attenuation.

Thus, we have

$$\begin{aligned}
\alpha^{OLS} &= \alpha\frac{\mathbb{V}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]} + \frac{\mathbb{E}\left[\Delta p^T_{a,n,t}\nu_{a,n,t}\right] + \mathbb{E}\left[\Delta p^F_{a,n,t}\nu_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]} \\
&= \alpha^{2SLS}\frac{\frac{\mathbb{V}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]}}{\frac{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]}{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]+\mathbb{E}\left[z_{a,n,t}\Delta p^F_{a,n,t}\right]}} + \frac{\mathbb{E}\left[\Delta p^T_{a,n,t}\nu_{a,n,t}\right] + \mathbb{E}\left[\Delta p^F_{a,n,t}\nu_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]}.
\end{aligned}$$

Thus, the regression (41) of α^{OLS}_h on α^{2SLS}_h for different horizons h yields:

$$\alpha^{OLS}_h = a + \theta\alpha_h + e_h,$$

[50]Here I assume that if the instrument for price $z_{a,n,t}$ correlates positively with $\Delta p^T_{a,n,t}$, it also correlates positively with $\Delta p^F_{a,n,t}$. That is, I make the reasonable assumption that this empirical noise trader demand shock does not have opposite impacts on the price changes analysts respond to and on those they do not respond to.

consistently estimates

$$\theta = \frac{\frac{\mathbb{V}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]}}{\frac{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]}{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]+\mathbb{E}\left[z_{a,n,t}\Delta p^F_{a,n,t}\right]}}.$$

As a result, the corrected covariance proportion from (42) is (again assuming all variables have been demeaned):

$$\begin{aligned}
\frac{\alpha_h^{2SLS}\mathbb{V}\left[\Delta p_{a,n,t}\right]}{Cov\left(\Delta p_{a,n,t},\Delta y^{(h)}_{a,n,t}\right)}\theta &= \alpha_h\frac{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]}{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]+\mathbb{E}\left[z_{a,n,t}\Delta p^F_{a,n,t}\right]}\frac{\mathbb{V}\left[\Delta p_{a,n,t}\right]}{Cov\left(\Delta p_{a,n,t},\Delta y^{(h)}_{a,n,t}\right)}\frac{\frac{\mathbb{V}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}\left[\Delta p_{a,n,t}\right]}}{\frac{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]}{\mathbb{E}\left[z_{a,n,t}\Delta p^T_{a,n,t}\right]+\mathbb{E}\left[z_{a,n,t}\Delta p^F_{a,n,t}\right]}}\\
&= \alpha_h\frac{\mathbb{V}\left[\Delta p^T_{a,n,t}\right]}{Cov\left(\Delta p_{a,n,t},\Delta y^{(h)}_{a,n,t}\right)},
\end{aligned}$$

as desired.