Journal of Financial Economics

The impact of prices on analyst cash flow expectations: Reconciling subjective beliefs data with rational discount rate variation ☆

Aditya Chaudhry

The Ohio State University, Fisher College of Business, United States of America

ARTICLE INFO

Keywords:
Asset pricing
Subjective beliefs
Expectations
Learning

ABSTRACT

I show that prices impact analyst cash flow expectations and argue this impact can partially reconcile subjective beliefs data with asset pricing models in which investors have rational expectations and discount rate variation drives prices. Previous work argues that correlations of biased *analyst* cash flow expectations with prices and future returns contradict rational models and imply biased *investor* expectations distort prices. However, using two instrumental variables for price, I find increases in price unrelated to cash flow news raise analyst cash flow expectations. Based on this empirical finding, I propose a model with rational investors that matches key moments in beliefs data: analysts form biased cash flow expectations by learning from prices that contain discount rate variation. Thus, while stylized facts in beliefs data can be consistent with investors having biased expectations that distort prices, these facts can also be consistent with investors having rational expectations and analysts learning from prices.

1. Introduction

Can models in which investors have rational expectations and discount rate variation drives asset prices resolve puzzles such as excess volatility and return predictability? A growing literature argues against this possibility because such models seem to conflict with stylized facts from subjective beliefs data. Specifically, equity research analyst cash-flow expectations not only correlate strongly with prices, they also have predictable forecast errors and negatively predict future returns.[1] Moreover, subjective expected returns correlate only weakly with prices.[2] By contrast, in models featuring investors with rational expectations, forecast errors are not predictable, cash flow expectations do not predict returns, and expected returns correlate strongly negatively with prices. Thus, previous work interprets these facts from subjective beliefs data as evidence that investors share analysts' biased cash flow expectations, which in turn impact prices (De Bondt and Thaler, 1990; Rafael, 1996; Bordalo et al., 2019, 2024; Delao and Myers, 2021, 2024; Nagel and Xu, 2021; Delao et al., 2023).

This paper, however, proposes a potential reconciliation between beliefs data and models featuring investors with rational expectations: I demonstrate *prices impact analyst cash flow expectations*. Using two complementary instrumental variables (IVs) for price from previous work, I find in the cross section of stocks that price increases unrelated to cash flow news raise analyst cash flow expectations. An exogenous 1% price increase driven by these IVs raises 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.

☆ Nikolai Roussanov was the editor for this article. I thank Niels Gormsen, Lars Hansen, Ralph Koijen, and Stefan Nagel for their invaluable advice and generous help. I thank Andrea Ricciardi for excellent research assistance. I thank Anna Pavlova and Taisiya Sikorskaya for providing access to data. For comments, I thank the editor, an anonymous referee, Alexandru Barbu, Francesca Bastianello, Filippo Cavaleri, Ricardo De La O, Zhiyu Fu, Xavier Gabaix, Andrei Goncalves, Valentin Haddad, Sebastian Hillenbrand, Zhiguo He, John Heaton, Sangmin Oh, Lubos Pastor, Yueran Ma, Sean Myers, Aaron Pancost, Anna Pavlova, Taisiya Sikorskaya, Alp Simsek, René Stulz, Pietro Veronesi, Michael Weisbach, Eric Zwick, and participants of the Chicago Booth Asset Pricing Working Group, Chicago Economic Dynamics Working Group, Chicago Booth Finance Brown Bag, Inter-Finance Ph.D. 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, the UW-Madison Junior Finance Conference, the 2024 Midwest Finance Association, the UVA McIntire Finance Seminar, and the ASU Carey Finance Brownbag. 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. This paper was previously part of my paper "Do Subjective Growth Expectations Matter for Asset Prices?".

[1] De Bondt and Thaler (1990), Rafael (1996), Chen et al. (2013), Delao and Myers (2021), Bordalo et al. (2019, 2024) and Delao et al. (2023)

[2] Delao et al. (2023) document this fact using analyst subjective expected returns in the cross section.

Received 5 May 2024; Received in revised form 10 May 2025; Accepted 14 May 2025

Available online 26 May 2025

This mechanism is economically significant: it explains about half of the covariance of prices with analyst expectations and forecast errors.

To illustrate how this mechanism can partially reconcile subjective beliefs data and models featuring investors with rational expectations, I propose such a model that quantitatively matches several subjective belief and cross-sectional asset pricing moments. Investors have private information (Grossman and Stiglitz, 1980; Hellwig, 1980) and discount rate variation drives excess volatility and return predictability. This private information motivates analysts to learn from prices as a signal of future cash flows. However, analysts inadvertently learn from discount rate variation in prices as well. Hence, analysts form biased cash flow expectations that predict returns and return expectations that correlate weakly with prices. Thus, private information generates an impact of prices on analyst expectations, and creates belief heterogeneity that allows the model to match subjective belief moments even though investors are rational.

Overall, my results suggest that models in which investors have rational expectations and discount rate variation drives asset prices need not conflict with stylized facts documented in subjective beliefs data. There remains an open question of what factors drive the discount rate variation that impacts prices and, in turn, analyst expectations. In particular, my empirical results suggest investor frictions and noise flows, in addition to risks and preferences, may be important drivers of such variation.[3] At the same time, it remains possible that investors do have biased cash flow expectations, or that some investors overreact to prices as analysts do. Ultimately, while analyst expectations remain a useful tool to explore these possibilities, they alone are likely insufficient. Since the impact of prices on cash flow expectations arises naturally in models with heterogeneous beliefs, my results suggest heterogeneity may be an important feature of subjective beliefs data and analyst expectations may not align with those of investors. Thus, determining the impact of *investor* beliefs on asset prices will likely require direct measures of investor beliefs[4] or empirical strategies accounting for heterogeneity between analysts and investors.[5] Additionally, since my results are in the cross section of equities, there is an open question of how prices impact cash flow expectations for the aggregate market.

I start by explaining the challenge in measuring the impact of prices on analyst cash flow expectations (in Section 2). If analysts and investors learn from the same cash flow news, then prices may *correlate* with analyst expectations without *impacting* them due to omitted variable bias. Thus, while some prior work uses reduced-form regressions to suggest prices impact analyst expectations, it neither quantifies this impact nor explores its relevance for asset pricing models.[6] To quantify this impact, I need instrumental variables that provide exogenous price variation: variation from noise trading unrelated to cash flow news. Section 2 presents my empirical framework and the relevance and exogeneity conditions for a valid instrument.

I use two instruments for price from previous work to overcome this challenge: changes in benchmarking intensity around Russell index reconstitutions (Pavlova and Sikorskaya, 2023) and mutual fund flow-induced trading (Lou, 2012; Li, 2022; Ben-David et al., 2022; Li et al., 2022; Van der Beck, 2021, 2022). Though these instruments use different assumptions, they yield consistent estimates of the impact of prices on analyst cash flow expectations, which underscores the robustness of this finding.

My first instrument for price is benchmarking intensity (BMI) changes around Russell index reconstitutions, following Pavlova and Sikorskaya (2023) (in Section 4). This instrument uses changes in the amount of benchmarked institutional capital following different stocks from funds tracking the Russell indices. Each *June*, stocks mechanically enter and exit the Russell 1000 and 2000 based on which side of a cutoff their *May* market caps fell on. Thus, the flows from benchmarked funds prompted by this reconstitution and the resulting price pressure are (conditionally) exogenous to June cash flow news (Chang et al., 2014; Crane et al., 2016; Glossner, 2019). In particular, the BMI measure of Pavlova and Sikorskaya (2023) captures the total change in benchmarked capital when a stock mechanically switches not only between the Russell 1000 and 2000 Blend indices, but also between the Value and Growth indices. Using June BMI changes for stocks near the market cap cutoffs to instrument for price, I find an exogenous 1% price increase raises analyst one to four year EPS expectations by 40 basis points.

My baseline specification using the BMI instrument addresses potential threats to instrument validity. First, I use May, not June, market caps to calculate the Russell cutoffs to avoid selection bias (Chang et al., 2014; Appel et al., 2021; Wei and Young, 2021). Second, I use the method of Ben-David et al. (2019) to approximate proprietary Russell market caps using standard databases, which allows accurate prediction of index assignment and avoids the threat posed by market cap mismeasurement (Glossner, 2019; Wei and Young, 2021).

In robustness checks I address additional potential threats (in Section 4.2). First, I show BMI increases do not forecast future profitability increases (consistent with Pavlova and Sikorskaya, 2023), which suggests analysts do not respond to real effects induced by BMI increases.[7] Second, I show the price impact of BMI changes reverts over time, as often expected of non-fundamental price changes.

My second instrument for price is the mutual fund flow-induced trading (FIT) instrument of Lou (2012), similar to the flow-to-stock instrument of Wardlaw (2020) (in Section 5). Flows induce funds to do some mechanical rebalancing: funds tend to scale their preexisting holdings proportionally in response to flows. This mechanical component of the cross-sectional trading induced by flows is uninformed and can provide exogenous price variation. An exogenous 1% price increase raises analyst one to four-year EPS expectations by 20 basis points (not statistically distinct from the estimate using the BMI instrument). This impact does not shrink as the forecast horizon grows. Moreover, the greater coverage of the FIT instrument than the BMI instrument (all stocks held by mutual funds versus only those in narrow windows around Russell cutoffs) 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, but do so over longer horizons as the price impact of the FIT instrument reverts.

My baseline specification using the FIT instrument addresses potential threats to instrument validity. First, I use the Lou (2012) FIT instrument, which is not subject to the Wardlaw (2020) critique of the Edmans et al. (2012) version of this instrument.[8] Second, I note that, as a shift-share instrument, FIT does *not* require exogenous flows (Goldsmith-Pinkham et al., 2020): correlations of flows with various

[3] For example, it is possible that investors have rational expectations and noise flows drive discount rate variation (e.g. as in Gabaix and Koijen, 2020). Such a model would seem to conflict with subjective beliefs data: investors' cash flow expectations would neither feature predictable forecast errors nor predict returns, and their expected returns would correlate strongly negatively with prices. The impact of prices on analyst cash flow expectations that I document offers a potential reconciliation between subjective beliefs data and this type of model, as well as other models with different sources of discount rate variation.

[4] E.g. Giglio et al. (2021), Dahlquist and Ibert (2024) and Couts et al. (2023)

[5] E.g. McCarthy and Hillenbrand (2021), Bianchi et al. (2023) and Chaudhry (2024)

[6] E.g. Brown et al. (1987), Lys and Sohn (1990), Abarbanell (1991), Forbes and Skerratt (1992), Guay et al. (2011) and Miller and Sedor (2014)

[7] E.g. Appel et al. (2016), Crane et al. (2016), Schmidt and Fahlenbrach (2017), Appel et al. (2019, 2021) and Heath et al. (2022)

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

objects[9] do *not* threaten exogeneity. The instrument's variation comes from variation in fund ownership shares across stocks, so a sufficient condition for exogeneity is that these ownership shares do not correlate with analyst belief shocks. Third, I show the price impact of FIT reverts over time, as often expected of non-fundamental price changes (consistent with Lou, 2012).

In robustness checks I address additional potential threats (in Section 5.3). First, I control for observed and latent stock characteristics (interacted with time fixed effects) associated with common fund styles to address the potential threat of ownership shares and analyst belief shocks depending on common characteristics. Second, I reconstruct the FIT instrument from only diversified funds to address potential threats posed by concentrated portfolios. Third, I reconstruct the instrument from only passive funds, which generally adhere to the proportional trading assumption, to address the potential threat of selection bias arising from systematic deviations from proportional trading (Berger, 2023). Fourth, I show FIT does not forecast future profitability increases, which suggests analysts do not respond to real effects of this instrument.

I next quantify how economically large the impact of prices on analyst cash flow expectations is (in Section 6). Specifically, I measure the proportions of the cross-sectional covariances of prices with analyst expectations and forecast errors explained by this channel. Doing so requires an estimate of the impact of average price changes on analyst expectations (the "average treatment effect"). Since the above estimates measure the impact of price changes driven by the two instruments on analyst expectations ("local average treatment effects"), they may differ from the impact of average price changes if analysts respond heterogeneously to different price changes (Angrist and Imbens, 1995). Thus, I use the method of Pancost and Schaller (2024) to recover (under certain assumptions discussed in Section 6.2) estimates of the impact of average price changes from these local estimates. I find the impact of prices on analyst cash flow expectations explains 60% of the covariance of prices with LTG expectations and 40% of the covariance with one-to-four year EPS expectations and forecast errors.

To illustrate how this impact of prices on analyst cash flow expectations can partially reconcile subjective beliefs data and models featuring investors with rational expectations, I propose such a model that matches several subjective belief and cross-sectional asset pricing moments (in Section 7). In the model, investors have rational cash flow expectations and private information; discount rate variation drives excess volatility and return predictability across stocks. Analysts seek to forecast cash flows, and so try to recover investors' private information by learning from prices (Grossman and Stiglitz, 1980; Hellwig, 1980). Yet since prices reflect discount rate variation, learning from prices introduces discount rate variation into analyst cash flow expectations. As a result, analyst cash flow expectations predict future returns. By inadvertently attributing discount-rate driven price variation to cash flow news, analysts also form subjective expected returns that correlate weakly with prices. Lastly, analysts overreact to prices (as in Glaeser and Nathanson, 2017; Bastianello and Fontanier, 2021, 2024; Bordalo et al., 2021), which creates predictable forecast errors. I estimate the model to match the impact of prices on analyst annual EPS expectations and the proportion of the covariance of analyst forecast errors with prices (40%) that this impact accounts for. I find the model quantitatively matches the predictability of returns by analyst cash flow expectations and the weak correlation of subjective expected returns with prices documented in previous work. I discuss extensions of this model with analyst overreaction to fundamental signals that can potentially account for the remaining 60% of the covariance of analyst forecast errors with prices, including extensions in which investors both do and do not remain rational.

This paper proceeds as follows. Section 2 introduces my empirical framework for measuring the impact of prices on analyst cash flow expectations. Section 3 discusses the data. Sections 4 and 5 measure the impact of prices on analyst expectations using BMI changes around Russell index reconstitutions and the FIT instrument. Section 6 presents the decomposition of the covariance of prices with analyst expectations. Section 7 presents the asset pricing model with rational investors that uses this impact to match asset pricing and subjective belief moments. Section 8 concludes.

1.1. Related literature

This paper relates to three bodies of literature: empirical work that uses analyst expectations to inform asset pricing models with subjective 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, a growing literature uses analyst cash flow expectations to test asset pricing models with subjective beliefs. Work going back at least to Malkiel (1970) uses analyst cash flow expectations as a proxy for investor expectations.[10] Under this assumption, previous work finds variation in subjective cash flow expectations can explain much (or all) of the time series and cross-sectional variation in stock prices.[11] Moreover, this literature finds analyst cash flow expectations negatively predict future returns[12] and feature predictable forecast errors.[13] Additionally, subjective expected returns correlate weakly with prices.[14]

These results seemingly contradict models with rational expectations in which forecast errors are not predictable, cash flow expectations do not predict returns, and expected returns correlate strongly negatively with prices. Thus, these results motivate models in which investors share biased analyst cash flow expectations, which distort prices (Bordalo et al., 2019, 2024; Nagel and Xu, 2021; Delao and Myers, 2021, 2024; Delao et al., 2023). In these models, prices do not impact cash flow expectations; biases arise only from biased learning from exogenous fundamentals (Bastianello and Fontanier, 2024).

This paper provides a potential reconciliation between the stylized facts from analyst expectations data and models featuring investors with rational expectations: I demonstrate that prices impact analyst cash flow expectations. While some previous work uses reduced-form regressions to suggest prices may impact analyst expectations, previous work does not quantify this impact nor determine if this impact is large enough to inform asset pricing models.[15] This is the first paper to quantify this impact and show it can potentially reconcile subjective beliefs data and models featuring investors with rational expectations.

Second, this paper relates to theoretical work on the impact of prices on cash flow expectations. My results are consistent with models featuring learning from prices due to dispersed information (e.g. Grossman and Stiglitz, 1980; Hellwig, 1980; Dubey et al., 1987; Kyle, 1989; Jackson, 1991; Mendel and Shleifer, 2012) and behavioral mislearning from prices (Glaeser and Nathanson, 2017; Bastianello and Fontanier,

[9] E.g. 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).

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

[11] E.g. Chen et al. (2013), Delao and Myers (2021, 2024) and Delao et al. (2023)

[12] E.g. Rafael (1996), Bordalo et al. (2019, 2024) and Nagel and Xu (2021)

[13] E.g. De Bondt and Thaler (1990), Rafael (1996), Bordalo et al. (2019) and Bordalo et al. (2024)

[14] In the cross section (Delao et al., 2023). The time-series evidence is more mixed depending on the agent surveyed (Wu, 2018; Delao and Myers, 2021; Nagel and Xu, 2023; Adam and Nagel, 2023).

[15] E.g. Brown et al. (1987), Lys and Sohn (1990), Abarbanell (1991), Forbes and Skerratt (1992), Guay et al. (2011) and Miller and Sedor (2014). Sulaeman and Wei (2019) find prices impact some analysts' buy recommendations.

2024, 2021; Bordalo et al., 2021). However, previous work using analyst expectations to test asset pricing models with subjective beliefs has generally not accounted for this learning from prices mechanism.[16] This paper not only demonstrates that analyst cash flow expectations behave in a manner consistent with learning from prices, it also quantifies the strength of this mechanism. Moreover, this is the first paper to show this mechanism is strong enough to potentially reconcile subjective beliefs data and models featuring investors with rational expectations. In particular, this paper provides one such quantitative model that matches several subjective belief and cross-sectional asset pricing moments.

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 corporate finance outcomes.[17] Other work uses Russell index reconstitutions to study the impact of institutional and passive ownership on corporate governance and product market outcomes.[18] This paper uses both instruments to study how exogenous price changes impact analyst cash flow expectations. This is the first paper to quantify this impact and show it is potentially large enough to reconcile subjective beliefs data and models featuring investors with rational expectations. I verify that real effects of both instruments do not drive my results.

2. Measuring the impact of prices on analyst expectations

The primary challenge in measuring the impact of prices on analyst cash flow expectations is omitted variable bias created by common information or sentiment shocks to analyst and investor expectations.

Consider this system of equations (microfounded in Appendix A with a model with private information):

$$\Delta p = Mz + \epsilon \tag{1}$$

$$\Delta y = \alpha \Delta p + \nu \tag{2}$$

Δy is the quarterly change in analyst expectations and Δp is the contemporaneous percentage price change (ex-dividend return). ϵ and ν are correlated and capture other cash flow news or sentiment shocks that both investors and analysts learn from, such as public signals like EPS announcements. z is a noise-trader demand shock that impacts price and is uncorrelated with these other cash flow news or sentiment shocks.

α is the impact of prices on analyst cash flow expectations that I measure. A regression of analyst expectations on prices does not identify α because if investors and analysts learn from common public signals, then ϵ and ν are positively correlated, which creates positive omitted variable bias (i.e. $\mathbb{E}[\Delta p \cdot \nu] \neq 0$).

The noise-trader shock z enables identification of α by providing exogenous price variation uncorrelated with ν. This two-stage least squares (2SLS) regression identifies α:

$$\Delta p = bz + e_1$$

$$\Delta y = \alpha \Delta \hat{p} + e_2.$$

To obtain a consistent estimate of α, the instrumental variable for price z must satisfy:

1. (Relevance) $M \neq 0$ in (1): The instrument has an effect on price.
2. (Exogeneity) $\mathbb{E}[z \cdot \nu] = 0$: The instrument affects analyst expectations only through price; it does not correlate with other shocks to analyst expectations, such as public signals that investors and analysts learn from. Thus, z must provide variation in prices unrelated to cash flow news.

To assess the economic significance of α, I calculate the proportion of the cross-sectional covariance of prices with analyst cash flow expectations explained by this impact of prices on analyst expectations:

$$\frac{\alpha \mathbb{V}[\Delta p]}{Cov(\Delta p, \Delta y)} = \frac{\text{Two Stage Least Squares Estimate of } \alpha}{\text{OLS Coefficient in Regression of } \Delta y \text{ on } \Delta \text{p}}. \tag{3}$$

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

Note that while, for simplicity, (2) assumes all price changes have the same impact on analyst cash flow expectations, in general different price changes may have heterogeneous impacts. In this case, 2SLS estimates of α are local average treatment effects (LATEs) (Angrist and Imbens, 1995) that reflect how price changes due specifically to the noise shocks z impact analyst expectations. These LATEs may differ from the average treatment effect (ATE): the impact of average price changes on analyst expectations. Appendix B derives conditions under which the LATE under- or overestimates the ATE.

Such heterogeneity does not alter the *qualitative* interpretation of 2SLS estimates of α: positive estimates imply prices do impact analyst cash flow expectations. However, it may alter the *quantitative* interpretation, which depends on how average price changes impact analyst expectations. Thus, to assess the economic significance of the impact of prices on analyst expectations I rely *not* on the 2SLS estimates of this impact, but instead on the proportion of the covariance between these objects that this impact explains. (3) correctly estimates this covariance decomposition when the LATE estimated by 2SLS equals the ATE, and Section 6 describes an adjustment I make using the method of Pancost and Schaller (2024) to correctly estimate this decomposition under general assumptions about heterogeneity.

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 EPS growth (LTG) expectations focused on by Bordalo et al. (2019, 2024) and Nagel and Xu (2021). I/B/E/S defines 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). Some work argues these expectations capture longer horizons of up to five to ten years (Sharpe, 2005). I/B/E/S reports LTG expectations at the stock × analyst institution × analyst × 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 × analyst institution × quarter panel of quarterly changes in LTG expectations.[19]

Second, I use the annual EPS expectations over shorter horizons of one to four years focused on by Delao and Myers (2021, 2024), and Delao et al. (2023). I/B/E/S reports EPS forecasts at the stock × fiscal year horizon × analyst institution × analyst × quarter level.

[16] Certain models of price extrapolation feature feedback from prices to cash flow expectations (e.g. Jin and Sui, 2022).

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

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

[19] Using institution-level (versus analyst-level) variation creates more quarter over quarter matches when computing quarterly expectations changes, which I winsorize at 5% to remove some extremely large outliers. Usually one analyst at an institution covers stock n in quarter t; if multiple analysts do, 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 inter-announcement price change ($\Delta p_{a,n,t}$) between the first such report in quarter $t-1$ and the last report in quarter t.

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

$$\Delta E_{a,n,t+4h|t+1} \equiv \frac{E_{a,n,t+4h|t+1} - E_{a,n,t+4h|t}}{E_{a,n,t+4h|t}}. \quad (4)$$

I drop all observations where $E_{a,n,t+4h|t} \leq 0$. Thus, I obtain a stock × analyst institution × quarter panel of quarterly revisions in h-year EPS expectations, where h ranges from one to four years.[21]

Each type of expectation has its advantages. Present value identities suggest the LTG expectations should correlate more strongly with prices, as documented in previous work (Bordalo et al., 2019, 2024, 2021). Yet 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 (4). Hence, the annual expectation revisions demonstrate that analyst forecast errors (in addition to cash flow expectations) rise in response to exogenous price increases (in magnitude, i.e. become more negative). Thus, analysts do not raise expectations due to exogenous price increases solely because these increases raise actual future earnings.

For both expectations types, multiple analyst institutions a (hereafter called analysts) report expectations for the same stock n in the same quarter t, often on different days. Thus, I let $\Delta y_{a,n,t}$ denote the quarterly change in a's expectations for stock n from quarter $t-1$ to t. I let $\Delta p_{a,n,t}$ denote 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 (which can differ across analysts for the same stock, as in Fig. 1).

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.[22]

[20] Delao and Myers (2021) follow the same interpolation procedure.

[21] I winsorize these final values at the 5% level (within each horizon h) to remove some extremely large outliers.

[22] Following Wardlaw (2020), I drop sector mutual funds when constructing the flow-induced trading instrument.

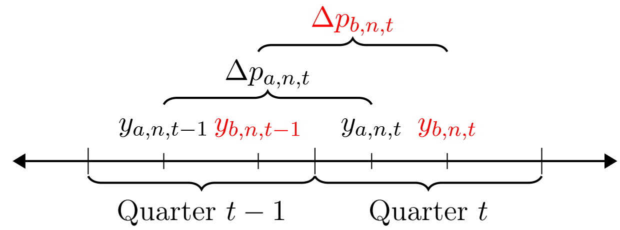


Fig. 1. Staggered timing of analyst reports.
Illustration of staggered timing of expectation releases for two analysts a and b for the same stock n.

4. Evidence from index reconstitutions

This section measures the impact of prices on analyst cash flow expectations using my first instrument for price: changes in benchmarking intensity around annual June Russell index reconstitutions.

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 June cash-flow news (Chang et al., 2014; Crane et al., 2016; Glossner, 2019). Thus, the June returns induced by this index reconstitution are also exogenous to June cash-flow 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.[23]

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

$$BMI_{n,t} = \sum_{\text{Index } j} \frac{\text{Institutional AUM Benchmarked to Index } j \text{ in Month } t \cdot \text{Weight of Stock } n \text{ in Index } j \text{ in Month } t}{\text{Stock } n \text{ Market Value in Month } t}.$$

$BMI_{n,t}$ 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 stock n 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 Russell reconstitution thresholds as an instrument for price. Stocks with larger ΔBMI experience more benchmarking inflows and more price pressure. While BMI is generally endogenous

[23] Technically all stocks in the Blend indices are in both the Value and Growth indices, just in different proportions.

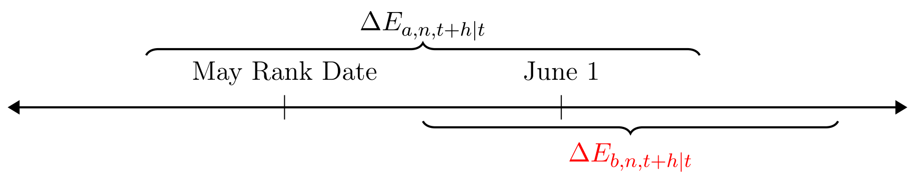


Fig. 2. Russell reconstitution timeline.
Illustration of Russell index reconstitution timing. $\Delta E_{a,n,t+h|t}$ is an h-year EPS expectation change for analyst a and stock n, where the original expectation was reported prior to the May rank date and the revised expectation is reported after June 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 4.2.2).

because index membership is, June BMI *changes* for stocks in this window are driven by Russell index membership changes, which are exogenous to June cash flow news conditional on the May rank-date market cap. Thus, $\Delta BMI_{n,t}$ satisfies the exogeneity condition and is uncorrelated with analyst belief shocks (conditional on controls $\boldsymbol{X}_{n,t}$ discussed below):

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

Controlling for the rank-date market cap removes the threat that bad news before the rank date could impact analyst beliefs (directly) and $\Delta BMI_{n,t}$ (by lowering market cap and moving stock n from the Russell 1000 to 2000). 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.

Section 4.2 addresses additional potential threats to exogeneity. Among other tests, I show long-run reversal of the price impact of ΔBMI and a lack of impact of ΔBMI on realized future earnings.

I run the following two-stage least squares regression:

$$\begin{aligned} \Delta p_{a,n,t} &= a_1 \Delta BMI_{n,t} + \beta_1' \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t} \\ \Delta E_{a,n,t+4h|t} &= \alpha \Delta \hat{p}_{a,n,t} + \beta_2' \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t,h}. \end{aligned} \quad (5)$$

The first stage regresses quarterly percentage price changes between analyst reports on June BMI changes.[24] The second stage regresses quarterly revisions to annual EPS expectations on instrumented price changes. $\boldsymbol{X}_{n,t}$ are stock-level controls discussed below and FE_t are year fixed effects.

I restrict the sample to analyst expectations changes that are exposed to June index reconstitutions: those for which the original expectation is reported in June or earlier and the revised expectation in June or later, as in Fig. 2. 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 to maintain comparability across firms (as in Pavlova and Sikorskaya, 2023). I use a 150-stock bandwidth in the baseline analysis (Section 4.2.1 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 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 D.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 et al. (2019) to approximate this proprietary market cap using standard databases.[25] Doing so predicts assignment to the Russell 1000 and 2000 with high accuracy, as shown in Appendix Table D2. Following previous work, I use May — not June — market caps to calculate the Russell reconstitution thresholds to avoid selection bias.[26] In this restricted sample, there is enough power to quantify the impact of prices on annual EPS expectations, but not on LTG expectations, which far fewer analysts cover (as detailed in Appendix Table D4).[27] I quantify the impact of prices on LTG expectations in Section 5 using the flow-induced trading instrument.

$\boldsymbol{X}_{n,t}$ includes stock-level controls used by Pavlova and Sikorskaya (2023): May rank-date log market cap, one-year monthly average bid–ask percentage spread[28], and the banding controls from Appel et al. (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. Conditional on these variables that determine Russell 1000/2000 membership, $\Delta BMI_{n,t}$ in June is exogenous.

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.[29]

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 (Delao and Myers, 2021, 2024).[30] The first stage regression in column 2 is strong: Russell reconstitution-driven BMI increases raise prices. The partial F-statistic (11.7) is

[24] Following Pavlova and Sikorskaya (2023), I winsorize price changes at 1%. Section 4.2 finds similar results at 0%.

[25] See Appendix Table D2 for details.

[26] E.g. Chang et al. (2014), Appel et al. (2021) and Wei and Young (2021)

[27] 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 D4 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.

[28] 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.

[29] 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 D1 for details.

[30] See Appendix C for a comparison of this OLS regression of analyst expectations on prices in changes to the regression in levels performed in previous work.

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 in each year 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.

Table 2
Impact 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	164 512	164 512	164 512	164 512
F	568.6	11.68	10.48	13.95
R-Squared	0.0825	0.00747	0.00152	

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 in the 150-stock window:

$$\Delta p_{a,n,t} = a_0 + a_1 \Delta BMI_{n,t} + \beta_1' X_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta E_{a,n,t+4h|t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \beta_2' X_{n,t} + FE_t + e_{2,a,n,t,h},$$

The first stage regresses percent price changes between analyst reports ($\Delta p_{a,n,t}$) on the June change in BMI ($\Delta BMI_{n,t}$) (column 2 reports the first-stage partial F-statistic). 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 ($\Delta \hat{p}_{a,n,t}$). $X_{n,t}$ 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.

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 impact of prices on annual EPS expectation revisions: an 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 concerns about how instrument strength impacts second-stage inference. Appendix Figure D1 displays first-stage and reduced-form binscatter plots. Appendix Table D3 displays alternate specifications.

Thus, prices impact analyst cash flow 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. Hence, analysts do *not* 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 enabling greater investment (Bernanke and Gertler, 1986) or because a firm's managers, customers, or suppliers learn from private information in prices and adjust real decisions (Subrahmanyam and Titman, 2001; Edmans et al., 2015)).

4.2. Robustness and threats to instrument validity

Fig. 3 summarizes the robustness checks I conduct for the baseline results in Table 2.

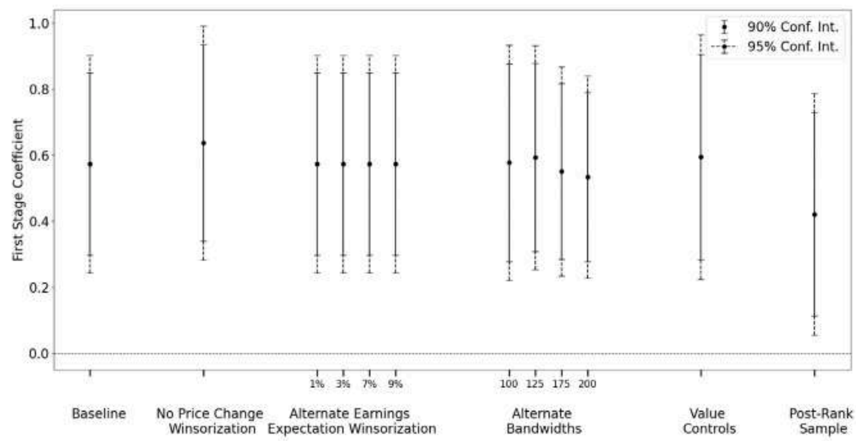
4.2.1. Alternative winsorizations and bandwidths

As displayed in Fig. 3, not winsorizing prices changes, instead of at the baseline 1%, yields similar first-stage (0.637 versus the baseline 0.573) and second-stage (α = 37.5 versus the baseline α = 41.6 basis points) estimates. Winsorizing EPS expectation revisions from 1% to 9%, instead of the baseline 5%, yields reduced-form estimates from 0.184 to 0.351 (statistically indistinct from the baseline 0.238), and second-stage α estimates from 32.1 to 61.2 basis points (statistically indistinct from the baseline 41.6 basis points).

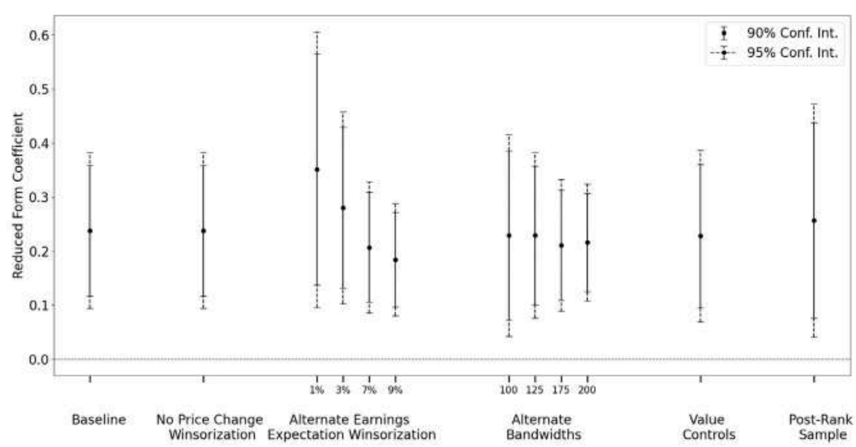
Using alternate bandwidths from 100 to 200 stocks around the cutoffs, instead of the baseline 150 stocks, yields similar first-stage (0.534 to 0.593 versus the baseline 0.573), reduced-form (0.211 to 0.229 versus the baseline 0.238), and second-stage (α = 38.2 to 40.2 versus the baseline 41.6 basis points)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

4.2.2. Real effects of BMI increases

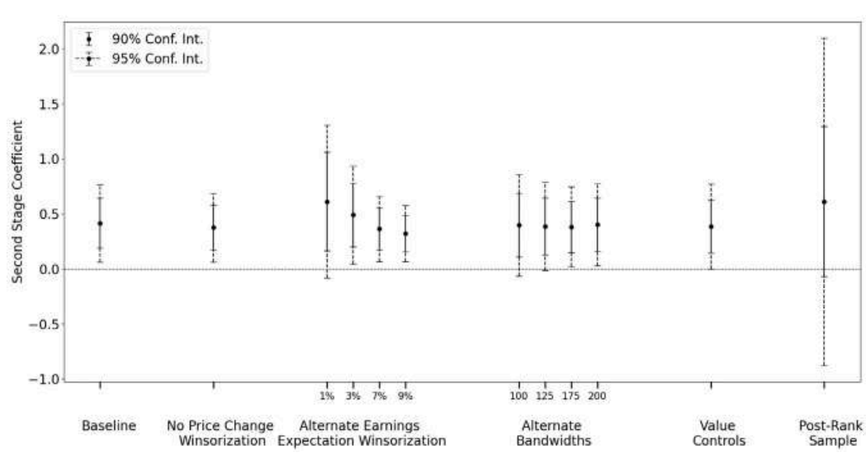
Appendix D.2.1 demonstrates BMI increases do not predict higher future EPS growth in levels or changes (as in Pavlova and Sikorskaya, 2023). These null results address the potential threat that BMI increases raise analyst expectations directly (instead of through prices) if



(a) First-Stage Results



(c) Reduced-Form Results



(c) Second-Stage Results

Fig. 3. Robustness checks for using ΔBMI as instrument.
This figure displays results for alternate specifications of two-stage least squares regression (5).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.

analysts pay attention to passive ownership increases and expect them to improve corporate governance or product market outcomes.[31]

I also repeat the baseline analysis using only analyst expectations changes for which the original expectation is reported *after the May rank date* but before the end of June (as in Fig. 2) and find similar results (displayed in Fig. 3), which further assuages this potential concern. If analysts do not respond to prices but expect passive ownership increases to raise future cash flows, then these anticipated improvements should appear in the first expectation reported after the May rank date (e.g. on June 1) because Russell index assignment depends only on information available at the rank date and can be accurately predicted with public data (see Appendix Table D2). 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 paying attention to passive ownership, but is consistent with analysts responding to prices.

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 is larger than, but statistically indistinct from, the baseline estimate ($\alpha = 61.1$ versus the baseline 41.6 basis points). Although this post-rank second-stage estimate is not statistically significant because the post-rank sample is ten times smaller than the full sample (11,980 versus 164,512 observations), the first-stage and reduced-form estimates remain significant.

4.2.3. Ex-post price reversal

Appendix Table D5 shows the price impact of BMI increases reverts in the long run (two years following the reconstitution), as is often expected of non-fundamental price changes.

4.2.4. Switchers between value and growth indices

I control for variables proxying for firm fundamentals and find similar results (displayed in Fig. 3). This test addresses the potential threat 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. Assignment to the Value and Growth indices at each May rank date is based on a custom algorithm applied to a proprietary database of ex-ante analyst forecasts, book-to-price ratio, and sales growth. Lacking these proprietary valuation metrics, I control for market-to-book ratio[32] 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 (α = 38.4 versus the baseline 41.6 basis points) estimates.

The post rank-date analysis discussed in Section 4.2.2 further assuages this concern. Analyst expectations changes after the rank date are not exposed to news from before the rank date that may impact assignment to the Value and Growth indices, but are still exposed to June reconstitution price changes.

5. Evidence from mutual fund flow-induced trading

This section measures the impact of prices on analyst cash flow expectations using my second instrument for price: mutual fund flow-induced trading (FIT).

Flows induce uninformed stock-level trading by mutual funds, which 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 cents to Apple, as Apple's weight in the S&P 500 is about 5%. This predicted mechanical component of the cross-sectional trading due to flows is uninformed.

I use the FIT instrument of Lou (2012) (similar to the flow-to-stock instrument of Wardlaw, 2020).[33] I first calculate the quarterly (percentage) flow to mutual fund i as

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

$TNA_{i,t}$ and $\text{Ret}_{i,t}$ are fund i's total net assets in quarter t and return from quarter $t-1$ to t. The predicted mechanical trading by fund i in stock n due to 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 expectations from $t-1$ to t. Aggregating across all funds and scaling by shares outstanding yields[34]:

$$\text{FIT}_{n,t} = \sum_{\text{fund } i} \underbrace{\frac{\text{Shares Held}_{i,n,t-2}}{\text{Shares Outstanding}_{n,t-2}}}_{\equiv S_{i,n,t-2}} f_{i,t}. \tag{6}$$

$S_{i,n,t-2}$ 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 F6 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 $v_{a,n,t}$ are *cross-sectionally* uncorrelated with $\text{FIT}_{n,t}$ for each analyst a and quarter t

$$\mathbb{E}\left[\text{FIT}_{n,t} v_{a,n,t}\right] = 0, \forall a, t, \tag{7}$$

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

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 FIT exogeneity (7) is that the ex-ante ownership shares are exogenous across stocks within each quarter t:

$$\mathbb{E}\left[S_{i,n,t-2} v_{a,n,t}\right] = 0, \forall a, i, t. \tag{8}$$

The sufficiency of cross-sectionally exogenous ownership shares follows from the result that exogenous shares are sufficient for shift-share instrument exogeneity (Goldsmith-Pinkham et al., 2020).

For example, if there is one fund, analyst, and quarter (drop subscripts i, a, and t), but N stocks, then $\text{FIT}_n = S_n f$ ($S_n \neq 1$ because there are other investors). FIT_n is exogenous if and only if the ownership shares are: $0 = \mathbb{E}\left[\text{FIT}_n v_n\right] = \mathbb{E}\left[S_n v_n\right] f$, because the flow is constant across stocks. Appendix E Proposition 2 (the same as Goldsmith-Pinkham et al., 2020 Proposition 2) generalizes this argument.

---

31 Previous work finds mixed results for the effect of passive ownership increases on corporate governance quality (e.g. Schmidt and Fahlenbrach, 2017; Appel et al., 2016, 2019, 2021; Heath et al., 2022). 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.

32 I construct book equity following the approach of Cohen et al. (2003).

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

34 Following Li (2022), I do not multiply the numerator by a "partial scaling factor" to adjust for funds scaling existing positions by less than one dollar per dollar of flow due to constraints. I use shares outstanding in the denominator (not shares held by all mutual funds as in Li, 2022) 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.

	$\Delta LTG_{a,n,t}$	$\Delta p_{a,n,t}$	$\mathrm{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num. Obs.	121 553.00	121 553.00	121 553.00	131 333.00
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

	$\Delta E_{a,n,t+4h\|t}$	$\Delta p_{a,n,t}$	$\mathrm{FIT}_{n,t}$	Num Stocks/(Fund, Quarter)
Num. Obs.	3396550.00	3396550.00	3396550.00	133 902.00
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 $\mathrm{FIT}_{n,t}$, 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.

Exogeneity of the ex-ante ownership shares (8) 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 exogenous flows. 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 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 FIT *cross-sectional* exogeneity. (7) can hold even if $\mathbb{E}\left[f_{i,t}\nu_{a,n,t}\right] \neq 0, \forall a, i, n$ because flows do not create cross-sectional variation in $\mathrm{FIT}_{n,t}$ across stocks within a quarter. Only the heterogeneous ownership shares create cross-sectional variation.

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 ((8) 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" can threaten exogeneity.

Section 5.3 explains the solution is to control for the problematic stock characteristics (size in this example) interacted with time fixed effects. Controlling for observed and latent characteristics that explain most ownership share variation yields similar results to the baseline analysis. Funds holding few stocks can raise similar issues, so I reconstruct the FIT instrument from only diversified mutual funds and find similar results. Systematic deviations from proportional trading can also raise similar issues (Berger, 2023), so I reconstruct the instrument from only passive funds and find similar results. I also show FIT does not forecast future profitability increases, suggesting analysts do not respond to real effects of this instrument.

5.1. Empirical results

I run the following two-stage least squares regression:

$$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + FE_t + e_{1,a,n,t}$$
$$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t}. \quad (9)$$

For analyst institution a and stock n in quarter t, $\Delta y_{a,n,t}$ is either the quarterly LTG expectation change $\Delta \mathrm{LTG}_{a,n,t}$, 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 a's two quarterly report dates for stock n in quarters $t-1$ and t. FE_t are quarter fixed effects.

Table 4 Panel (a) presents the LTG expectations results.[35] The OLS regression of LTG expectation changes on price changes in column 1 finds a strong association between these objects, as in previous work (Bordalo et al., 2019, 2024; Nagel and Xu, 2021).[36] 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 impact: 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. The smaller magnitude for this α estimate than for the annual EPS expectations is reasonable since the LTG expectations likely capture a more persistent component of cash flow expectations (as discussed in Section 7.1.2).

Panel (b) displays the one to four-year annual EPS expectations results. The OLS regression of EPS expectation revisions on price changes in column 1 finds a strong association between these objects, as in previous work (Delao and Myers, 2021, 2024). 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:

[35] Appendix Table F7 displays results with block-bootstrapped (instead of quarterly-clustered) confidence intervals.

[36] See Appendix C for a comparison of this OLS regression of analyst expectations on prices in changes to the regression in levels performed in previous work.

Table 4
Impact of Prices on Analyst Expectations Using FIT as Instrument.

Panel (a): LTG Expectation Changes	(1) OLS	(2) First Stage	(3) Reduced Form	(4) 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	121 553	121 553	121 553	121 553
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	3 396 550	3 396 550	3 396 550	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 ($\text{FIT}_{n,t}$) (column 2 reports the first-stage partial F-statistic). The reduced form 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 $\text{FIT}_{n,t}$. The second stage regresses $\Delta \text{LTG}_{a,n,t}$ in Panel (a) and $\Delta E_{a,n,t+4h|t}$ 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.

higher flow-induced trading raises annual EPS expectation revisions. The second-stage α estimate in column 4 reveals a statistically and economically significant impact: an exogenous 1% price increase raises annual EPS expectations by 21 basis points, which is statistically indistinct from the 41 basis points estimate using the ΔBMI instrument in Table 2.

Since these revisions of fixed-horizon EPS expectations reflect changes in forecast errors, as Section 3 discusses, this $\alpha > 0$ result implies forecast errors increase (in magnitude, i.e. become more negative) in response to exogenous price increases. Thus, this result is inconsistent with analysts raising their cash flow expectations due to exogenous price increases solely because these increases raise actual future cash flows.

Appendix Figure F5 displays first-stage and reduced-form binscatter plots. Figures F7 and F8 display alternate specifications. Figure F9 displays results for alternate winsorizations.

5.2. Persistence in the term structure and long-run reversal

The impact of prices on analyst expectations does not shrink as forecast horizon grows or over the following year for a given forecast horizon. This impact reverts in the long run as the price impact of FIT reverts.

Fig. 4 displays the results of two-stage least squares regression (9) 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 F6 displays first-stage and reduced-form results. Thus, the impact of prices on analyst cash flow expectations appears permanent in the term structure of expectations: it does not shrink as forecast horizon grows.

Do analysts revise their expectations ex-post once they have enough public information to realize they have responded to price changes 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 Fig. 1). Funds report these holdings to the SEC with a 60-day delay, so they become public in quarter $t-1$.[37] If 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 they later learn of these filings, construct $\text{FIT}_{n,t}$, realize they have responded to noise, and revise their expectations to remove this noise, then the impact of $\text{FIT}_{n,t}$ on analyst expectations should revert quickly.

These quick ex-post revisions do not occur over one year. I regress expectations changes on lagged FIT:

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

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

Fig. 5 displays the sum of the coefficients on the lagged FIT instruments for each maximum lag $L = 0, \dots, 4$: $\sum_{s=0}^{L} \beta_s$. 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 Fig. 1). At

[37] The Thomson Reuters S12 database collects mutual fund holdings from SEC Forms N-30D, N-Q, and N-CSR, and voluntary disclosures (Zhu, 2020). Form N-Q is filed within 60 days after a fund's 1st and 3rd fiscal quarters (Securities, U.S. and Exchange Commission, 2023b). 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... pursuant to Rule 30e-1 under the [Investment Company Act of 1940]" (Securities, U.S. and Exchange Commission, 2023a). These reports are made within 60 days after a fund's 2nd and 4th fiscal quarters (Office U.S. Government Publishing, 2023).

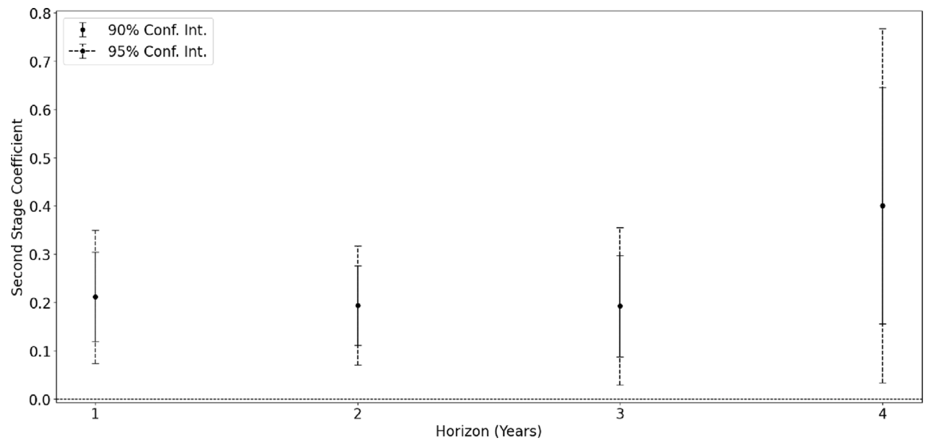


Fig. 4. Impact of prices on term structure of EPS expectations.
This figure displays the following two-stage least squares regressions results:
$\Delta p_{a,n,t} = a_0 + a_1 \text{FIT}_{n,t} + FE_t + e_{1,a,n,t}$
$\Delta E_{a,n,t+4h|t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + FE_t + e_{2,a,n,t}$,
The 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 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.

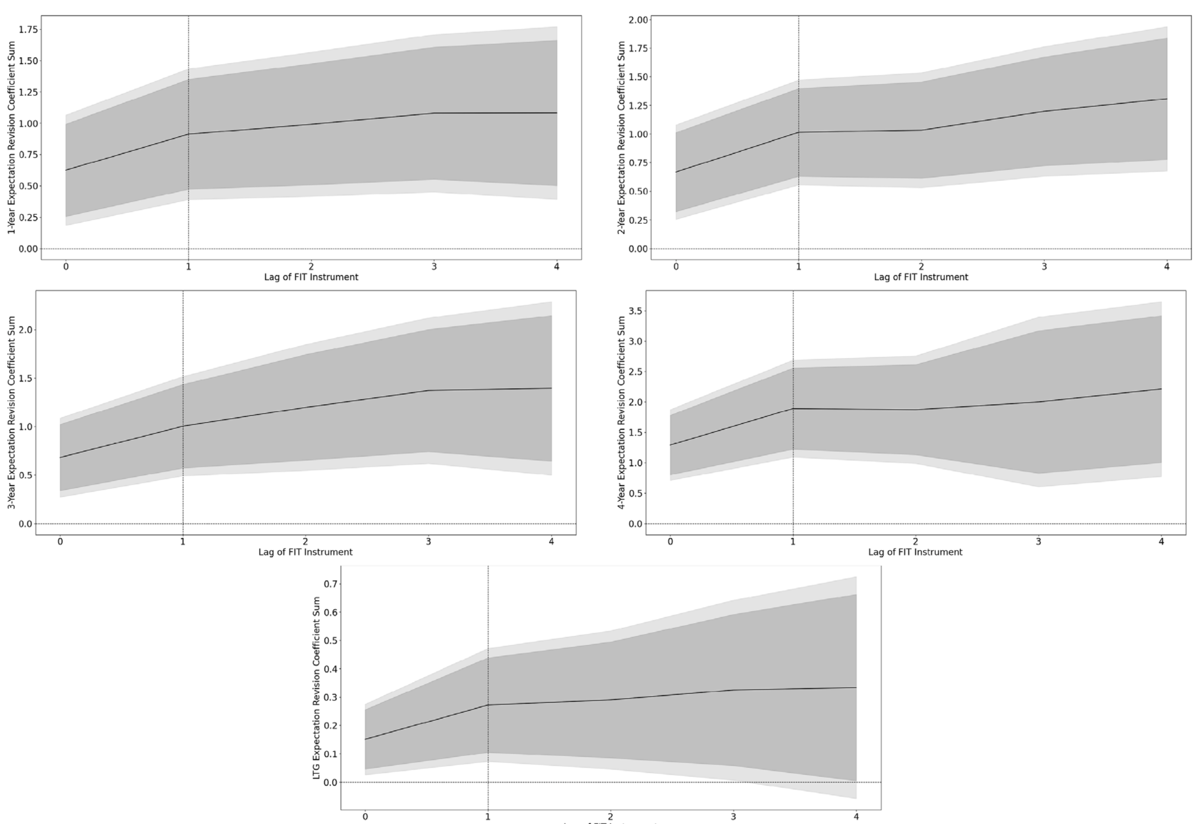


Fig. 5. Lack of immediate reversal in impact of FIT on analyst cash flow expectations.
This figure displays the coefficient sums $\sum_{s=0}^{L} \beta_s, L = 0, \ldots, 4$, from the following regression:

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

$\Delta y_{a,n,t}$ 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.

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 F10 shows this impact reverts at longer horizons (twelve to sixteen quarters) once FIT price impact reverts (Lou, 2012 finds reversal at similar horizons of eight to sixteen quarters). If prices impact analyst expectations, then analysts should lower their expectations due to ex-post price reversals. I find evidence of this prediction, but lack power at longer lags for a sharp statistical test.

5.3. Robustness and threats to instrument validity

This section presents robustness checks to address the main potential threat to FIT exogeneity: that both ex-ante ownership shares and analyst belief shocks depend on common stock characteristics. As discussed above, since ownership share exogeneity is sufficient for FIT exogeneity, exogeneity is *not* threatened by flows chasing returns and stock characteristics, nor by flows correlating with surveyed beliefs and past flows.

If ownership shares $S_{i,n,t-2}$ and analyst belief shocks $v_{a,n,t}$ depend on common stock characteristics, they correlate across stocks within a quarter and (8) fails: the ownership shares are not exogenous. 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 expectations more for small stocks than large stocks. Thus, stocks with higher small-cap fund ownership shares have higher analyst belief shocks, which violates (8).

To be concrete, consider this factor structure in ownership shares and analyst belief shocks[38]:

$$S_{i,n,t-2} = \boldsymbol{c}_i' \boldsymbol{X}_n + \tilde{S}_{i,n,t-2} \tag{10}$$

$$v_{a,n,t} = \boldsymbol{X}_n' \boldsymbol{\eta}_t + \tilde{v}_{a,n,t}. \tag{11}$$

In (10), 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 (11), 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 $\tilde{v}_{a,n,t}$ are uncorrelated with other objects.

In this case, the ownership shares are *not* cross-sectionally exogenous ((8) fails):

$$\forall a, i, t : \mathbb{E}\left[S_{i,n,t-2} v_{a,n,t}\right] = \boldsymbol{c}_i' \mathbb{E}\left[\boldsymbol{X}_n \boldsymbol{X}_n'\right] \boldsymbol{\eta}_t \neq 0. \tag{12}$$

Since the ownership shares are not cross-sectionally exogenous, neither is $\mathrm{FIT}_{n,t}$ ((7) fails):

$$\mathrm{FIT}_{n,t} = \sum_i f_{i,t} S_{i,n,t-2} = \underbrace{\left(\sum_i \boldsymbol{c}_i f_{i,t}\right)'}_{\equiv \boldsymbol{\beta}_t'} \boldsymbol{X}_n + \underbrace{\sum_i \tilde{S}_{i,n,t-2} f_{i,t}}_{\equiv \tilde{\mathrm{FIT}}_{n,t}} \tag{13}$$

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

Simply put, the problem here is that cross-sectional FIT variation comes from heterogeneous exposures to flows (which come from heterogeneous ownership shares) that correlate with exposures to other aggregate shocks: small stocks are more exposed to both small-cap fund flows and the tax news.

The solution is to control for characteristics $\boldsymbol{X}_n$ interacted with quarter indicators. Doing so removes the variation in $\mathrm{FIT}_{n,t}$ from these characteristics ($\boldsymbol{\beta}_t' \boldsymbol{X}_n$). From (13), using $\mathrm{FIT}_{n,t}$ while linearly controlling for characteristics interacted with quarter indicators is the same as constructing the instrument from residual ownership shares $\tilde{S}_{i,n,t-2}$: $\tilde{\mathrm{FIT}}_{n,t} = \sum_i \tilde{S}_{i,n,t-2} f_{i,t}$. $\tilde{\mathrm{FIT}}_{n,t}$ is exogenous because the residual shares are.[39]

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

[39] $\tilde{\mathrm{FIT}}_{n,t}$ is exogenous ($\mathbb{E}\left[\tilde{\mathrm{FIT}}_{n,t} v_{a,n,t}\right] = 0, \forall a, t$) since $\tilde{S}_{i,n,t-2}$ are exogenous ($\mathbb{E}\left[\tilde{S}_{i,n,t-2} v_{a,n,t}\right] = 0, \forall a, i, t$).

Which characteristics should one control for? While I cannot identify which characteristics correlate with unobserved analyst belief shocks, 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 5.3.1 controls for stock characteristics associated with the investment styles of funds that drive most FIT variation and that may threaten exogeneity. Since ownership shares and analyst belief shocks may also depend on unobserved characteristics, Section 5.3.2 controls for latent characteristics from a latent factor model that explain most cross-sectional ownership share variation. Section 5.3.3 reconstructs FIT from only diversified funds to address the similar issues raised by funds holding few stocks: ownership shares depend on a specific stock characteristic — firm identity. Section 5.3.4 reconstructs FIT from only passive funds, which generally adhere to proportional trading, to address similar issues raised by systematic deviations from proportional trading (Berger, 2023). Section 5.3.5 shows FIT does not predict future earnings growth to address the potential threat of analysts responding to real effects of this instrument.

As summarized in Fig. 6, these alternate specifications yield α estimates similar to the baseline estimates in Table 4. Appendix Figures F14 and F15 present the first-stage and reduced-form results.

5.3.1. Controlling for observed stock characteristics

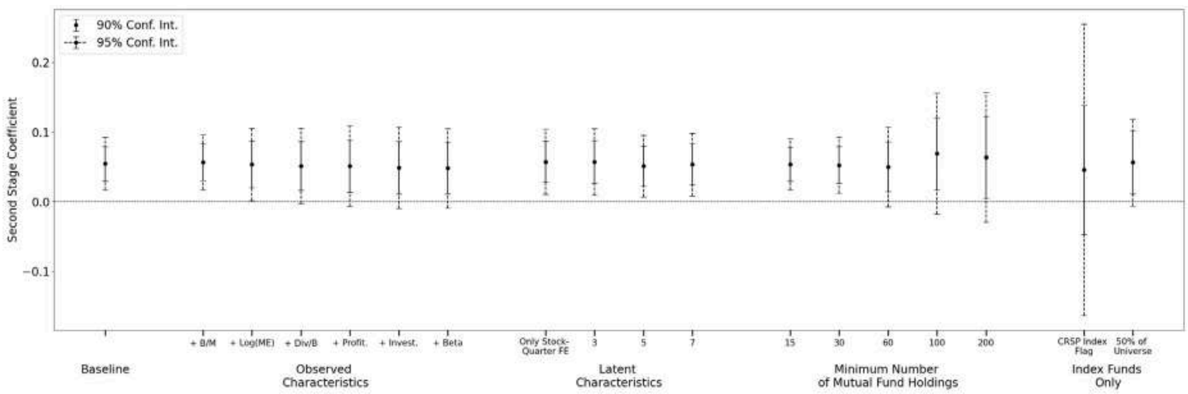
I control for stock characteristics associated with the fund styles that drive most of the variation in the FIT instrument. Since investment style is one of the most important determinants of ownership shares within a fund (e.g. small-cap funds have larger ownership shares in small stocks than in large stocks), style characteristics may threaten ownership share exogeneity and should be controlled for.

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 et al., 2020). One can compute which funds, and so which styles, receive more weight in this estimation using the "Rotemberg weights" from Goldsmith-Pinkham et al. (2020) (see Appendix F.5 for details). Violation of share exogeneity (8) for higher Rotemberg weight styles biases the α estimates more. Thus, characteristics associated with high-weight styles are the most important to control for.

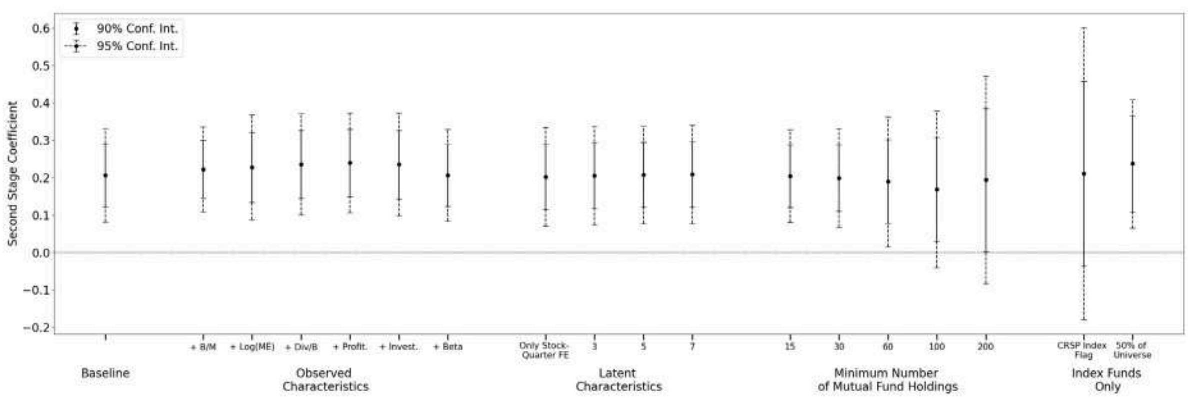
As displayed in Table 5, the most important styles (identified from CRSP style codes[40]) for both the LTG and annual EPS expectations samples are cap-based and growth/income-based. The top four styles account for 116% and 88% of the total Rotemberg weight for these samples, respectively, so these styles drive almost all FIT instrument variation (a few styles have slightly negative weights, as detailed in Appendix F.5). For example, since the most important style for the LTG sample is growth, much of 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). If analyst belief shocks for growth and value stocks having heterogeneous exposures to other aggregate shocks, that may violate share exogeneity. Thus, size and valuation metrics may threaten share exogeneity and should be controlled for because cap-based and growth/income-based fund ownership shares correlate with these characteristics.

I control for log market equity as a size measure, book-to-market and dividend-to-book ratios as valuation metrics, and profitability,

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



(a) LTG Expectations



(b) Annual EPS Expectations

Fig. 6. Robustness checks for using FIT instrument.
This figure displays the following two-stage least squares regression results:
$\Delta p_{a,n,t} = a_0 + a_1 \mathrm{FIT}_{n,t} + \boldsymbol{\beta}_1^{'} \boldsymbol{X}_{n,t} + FE_t + e_{1,a,n,t}$
$\Delta y_{a,n,t} = b_0 + \alpha \Delta \hat{p}_{a,n,t} + \boldsymbol{\beta}_2^{'} \boldsymbol{X}_{n,t} + FE_t + e_{2,a,n,t}.$
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 reconstructs $\mathrm{FIT}_{n,t}$ 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 reconstruct $\mathrm{FIT}_{n,t}$ 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 four 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.

investment, and market beta as popular characteristics that may correlate with ownership shares.[41] These six characteristics explain 46% of within fund-quarter ownership share variation (see Appendix F.6 for details).

As Fig. 6 displays, controlling for these characteristics interacted with quarter indicators does not affect the α estimates. Sequentially adding controls, the LTG expectations 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 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).

5.3.2. Controlling for latent stock characteristics

I fit a latent factor model to address the potential threat of ownership shares and analyst belief shocks depending on unobserved characteristics:

$$S_{i,n,t} = c'_{i,t} \boldsymbol{X}_{n,t} + FE_{i,t} + FE_{n,t} + \tilde{S}_{i,n,t}. \tag{14}$$

I fit factor model (14) to the fund × stock $S_{i,n,t}$ panel in each quarter using regularized singular value decomposition (Funk, 2006) (see Appendix F.7 for details).[42] The stock-quarter fixed effect and first seven characteristics explain 75% of within fund-quarter ownership share variation (see Appendix F.6 for details).

As Fig. 6 displays, controlling for these eight latent characteristics interacted with quarter indicators does not affect the α estimates. The LTG 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 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 F16, F17, and F18 display the first-stage, reduced-form, and second-stage results for different numbers of latent factors.

5.3.3. Requiring minimum number of holdings

I reconstruct the FIT instrument from only mutual funds with many holdings to address the potential threat posed by concentrated portfolios. Funds holding few stocks may violate share exogeneity. In the extreme case where each fund holds one stock, any stock-specific shocks (e.g. earnings surprises) will violate share exogeneity. For example, 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.[43] Since ownership shares and analyst belief shocks depend on a common characteristic, share exogeneity (8) is violated, as in (12).

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).

To further address this concern, I reconstruct $\text{FIT}_{n,t}$ using only funds with at least $M \in [15, 200]$ holdings in quarter t. As Fig. 6 displays, doing so does not affect the α estimates. The LTG expectations estimates in Panel (a) range from $\alpha = 5.0$ to 6.9 basis points, similar to the baseline $\alpha = 5.5$ basis points estimate in Table 4. The annual EPS expectations estimates in Panel (b) range from $\alpha = 16.8$ to 20.4 basis points, similar to the baseline $\alpha = 20.6$ basis points estimate. Power decreases as the minimum number of holdings rises because excluding some funds reduces the FIT instrument's variation.

5.3.4. Using only passive funds

I reconstruct the FIT instrument from only passive funds to address the potential threat posed by systematic deviations from the proportional trading assumption. The FIT instrument's construction assumes that funds tend to scale holdings proportionally in response to flows, or deviate from proportional trading at random. However, Berger (2023) finds systematic deviations from proportional trading that can create selection bias: 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). For the Edmans et al. (2012) $MFFLOW$ instrument, Berger (2023) argues these systematic deviations occur for stocks with certain characteristics (e.g. illiquid stocks) because this instrument uses only the predicted trading driven by extreme outflows.[44] 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 a dependence of ownership shares on stock characteristics, which threatens share exogeneity (8).[45]

Reconstructing the FIT instrument from only passive funds addresses this concern because Berger (2023) finds passive funds, which closely track an index, generally adhere to proportional trading, or at least deviate much less from it than active funds do.[46] Thus, if selection bias drives the baseline results in Table 4, then constructing the FIT instrument from only passive funds should yield smaller α estimates. Yet doing so yields similar results, which suggests this selection bias concern is not serious empirically.

I use two definitions of "passive funds". First, I classify funds with the CRSP index fund flag[47] or "target date" in their names as passive, as in Berger (2023). As Fig. 6 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, similar to the baseline 5.5 and 20.6 basis points estimates in Table 4. These point estimates are not statistically significant because the sample sizes are much smaller since the CRSP index fund flag is only available since 1998 and the first named "target date" funds appear in my sample in 2005.[48]

---

[41] 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. 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.

[42] 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.

[43] If there are N total stocks, $\boldsymbol{X}_n$ in (10) and (11) is an N-dimensional vector of stock indicators: $\boldsymbol{X}_n = \left[1_{j=n}\right]_{j=1}^{N}$. Firm identity cannot be controlled for: one would have to control for stock-quarter fixed effects, which would absorb all FIT variation.

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

[45] Let actual trading by fund i in stock n due to flow $f_{i,t}$ be $\psi_{i,n,t} \cdot f_{i,t}$. Define $\text{FIT}^{\psi}_{n,t} = \sum_i S^{\psi}_{i,n,t} f_{i,t}$, where $S^{\psi}_{i,n,t} = \psi_{i,n,t}/\text{Shares Outstanding}_{n,t}$. If deviations of actual trading from proportional trading depend on stock characteristics $\boldsymbol{X}_{n,t}$ ($\text{Shares Held}_{i,n,t-2} \cdot f_{i,t} = \psi_{i,n,t} \cdot f_{i,t} + c' \boldsymbol{X}_{n,t} \cdot f_{i,t}$), then deviations of the shares the FIT instrument uses ($S_{i,n,t}$) from the shares that "should" be used ($S^{\psi}_{i,n,t}$) depend on $\boldsymbol{X}_{n,t}$: $S_{i,n,t} = \left(S^{\psi}_{i,n,t} \cdot \text{Shares Outstanding}_{n,t} + c' \boldsymbol{X}_{n,t}\right) / \text{Shares Outstanding}_{n,t-2}$.

[46] Heath et al. (2022) note passive ETFs hold a (large) sample of stocks from their benchmarks that trades off tracking error versus transactions costs (they hold 97%, 89%, and 63% of stocks in the top, middle, and bottom liquidity terciles).

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

[48] 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.

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 a fund's investment universe 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.[49] For $M = 50\%$, this rule successfully classifies 84% of fund-quarter observations as passive or not based on the above CRSP index fund flag and "target date" name definition, with a false positive rate of only 8%. As Fig. 6 displays, reconstructing 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 accuracy and yields similar results, at the expense of power, as Appendix Figure F21 displays.[50] Appendix Figures F19 and F20 display the first-stage and reduced-form results.

5.3.5. Real effects of FIT

Appendix F.3 demonstrates FIT does not predict future EPS growth in levels or changes in this sample. These null results address the potential threat of FIT raising analyst expectations directly (instead of through prices) because analysts respond to real effects of the FIT instrument.[51]

6. Quantifying the importance of the impact of prices on analyst expectations

This section measures the proportion of the covariance between prices and analyst cash flow expectations explained by the impact of prices on analyst expectations. This covariance proportion requires an estimate of the impact of average price changes on analyst expectations (the "average treatment effect"). Since the two-stage least squares estimates in Sections 4.1 and 5.1 measure the impact of price changes specifically driven by the two instruments on analyst expectations ("local average treatment effects"), they may differ from the impact of average price changes if analysts respond heterogeneously to different price changes (Angrist and Imbens, 1995).

Section 6.1 describes how I use the approach of Pancost and Schaller (2024) to recover estimates of this average impact and presents the empirical estimates of the decomposition of the covariance of prices with analyst expectations. Section 6.2 provides sufficient conditions under which this approach recovers the true covariance decomposition in a general setting. Section 6.3 demonstrates that these covariance proportion estimates prove robust to relaxations of these conditions.

6.1. Empirical covariance decomposition

For now, consider a simple setting (that I generalize in Section 6.2) with two types of days: type-T days on which price changes impact analyst cash flow expectations, and type-F on which they do not. In this case we can write the quarterly changes in price $\Delta p_{a,n,t}$ and analyst expectations $\Delta y_{a,n,t}$ as

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

The true covariance proportion explained by the impact of prices on analyst expectations is

$$\frac{\alpha \mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]}{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y_{a,n,t}\right)} = \frac{\text{Two Stage Least Squares Estimate of } \alpha}{\text{OLS Coefficient in Regression of } \Delta y_{a,n,t} \text{ on } \Delta p_{a,n,t}} \cdot \underbrace{\frac{\mathbb{V}^{CX}\left[\Delta p^T_{a,n,t}\right]}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}}_{\equiv \theta \leq 1}, \tag{15}$$

where the superscript CX (for "cross-sectional") means time fixed effects are removed.

The proportion of total price variation that impacts analyst expectations is θ, and those price changes have an impact of α on analyst expectations. Thus, $\alpha^{ATE} \equiv \alpha \cdot \theta$ represents the impact of average price changes on analyst cash flow expectations (i.e. average treatment effect). As discussed in Section 6.2, in general the true covariance proportion explained is α^{ATE} divided by the OLS coefficient.

Thus, measuring the true covariance proportion (15) requires an estimate of θ in addition to the two-stage least squares and OLS estimates. I use the method of Pancost and Schaller (2024) to measure θ. Let α_h^{2SLS} 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" (due to $\Delta p_{a,n,t}$ being a "noisy proxy" for $\Delta p^T_{a,n,t}$ given the presence of $\Delta p^F_{a,n,t}$), then α_h^{OLS} is a linear function of α_h^{2SLS}:

$$\alpha_h^{OLS} = \theta \alpha_h^{2SLS} + OVB_h,$$

Pancost and Schaller (2024) demonstrate that in an OLS meta-regression of α_h^{OLS} on α_h^{2SLS}

$$\alpha_h^{OLS} = a + b^{meta} \alpha_h^{2SLS} + e_h, \tag{16}$$

b^{meta} consistently estimates θ. Hence, $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta}$ recovers corrected proportion (15).[52]

I use the horizon-specific α_h^{2SLS} and α_h^{OLS} estimates obtained from the FIT instrument to compute this meta-regression and corrected covariance proportion. The horizon-specific α_h^{2SLS} estimates obtained from the BMI instrument are noisy (see Appendix Table D4) due to the small sample used (stocks in a narrow window around the reconstitution thresholds in June), which leads *underestimation* of both b^{meta} and the true covariance proportion explained (Pancost and Schaller, 2024). Thus, I defer the covariance decomposition using the BMI instrument to Appendix G Table G9.

Fig. 7 displays the corrected covariance proportion (15) for the LTG and annual EPS expectations based on the α_h^{2SLS} and α_h^{OLS} estimates from the FIT instrument. I estimate $b^{meta} = 0.54$, which implies an average 1% price increase raises analyst LTG and one to four-year annual EPS expectations by $\alpha_h^{ATE} = \alpha_h^{2SLS} \cdot b^{meta} = 2.7$ and 11.1 basis points, respectively (see Appendix Table G10 for details). Accordingly, the impact of prices on analyst cash flow expectations explains roughly $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta} = 60\%$ (61.0%) and 40% (39.2%) of the cross-sectional covariances of prices with LTG and annual EPS expectations. Since the annual expectation revisions capture forecast error changes, this impact also explains roughly 40% of the cross-sectional covariance of prices with forecast errors. Common information or sentiment shocks to analysts and investors explain the remainder. Appendix G provides estimation details; Table G10 provides covariance decompositions and α_h^{ATE} estimates for all horizons, as well as confidence intervals for b^{meta}.

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

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

[51] These results are consistent with Wardlaw (2020), which shows that after correcting the mechanical endogeneity issue in the Edmans et al. (2012) version of the mutual fund flow-induced trading instrument used in much previous work, many of the real effects documented in previous work no longer hold. I use the Lou (2012) FIT instrument, which is not subject to the Wardlaw (2020) critique of the Edmans et al. (2012) version of this instrument.

[52] Appendix G.3 presents a simulation exercise to confirm that this meta-regression performs well in my small-sample setting (i.e. the number of observations in the meta-regression is the number of horizons, so five: $h = 1, 2, 3, 4$, and LTG).

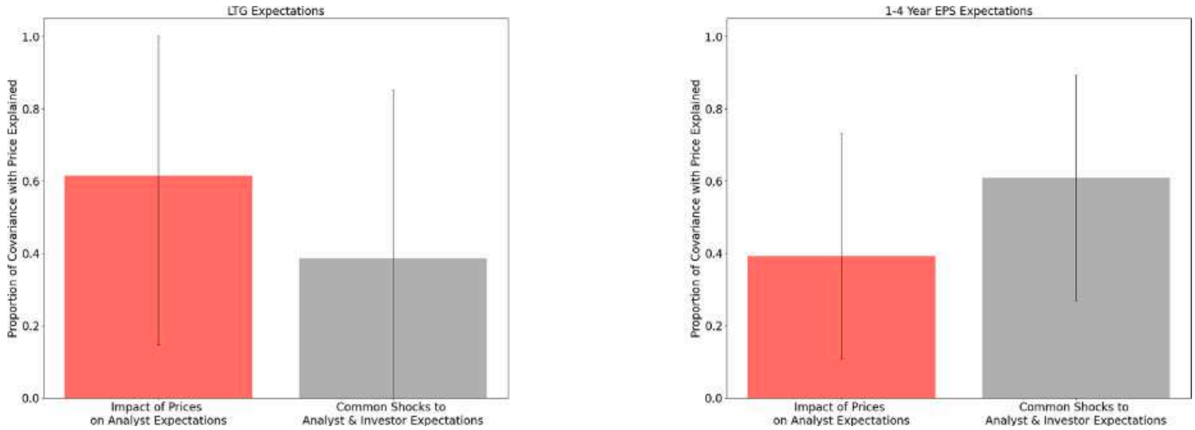


Fig. 7. Decomposition of covariance between analyst cash flow expectations and prices.
This figure displays the proportion of the covariance of analyst cash flow expectation changes with contemporaneous price changes explained by the impact of prices on analyst cash flow expectations (red) and by common information or sentiment shocks to analyst and investor expectations (gray). These proportions sum to one. The left and right panels display this decomposition for changes in LTG expectations and revisions to one to four year EPS expectations (which capture forecast error changes). Error bars represent quarterly block-bootstrapped 95% confidence intervals.

These adjusted covariance proportions are more conservative than the unadjusted proportions. The two-stage least squares and OLS estimates (41 and 26 basis points) from the BMI instrument in Table 2 imply the impact of prices on annual analyst EPS expectations explains 157% of the covariance between these objects (following (3)). The two-stage least squares (5 or 21 basis points) and OLS estimates (4 or 28 basis points) from the FIT instrument in Table 4 imply the impact of prices on analyst expectations explains 125% and 75% of the observed covariances of prices with the LTG and annual EPS expectations. Thus, the adjusted covariance proportions reported in Fig. 7 attribute less of this covariance to the impact of prices on analyst expectations than the unadjusted covariance proportions do.

6.2. *Sufficient conditions for accurate covariance decomposition*

In a general setting, under three assumptions, $\alpha_h^{2SLS} \cdot b^{meta}$ recovers α_h^{ATE} and $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta}$ recovers the true proportion of the covariance of prices with analyst expectations explained by the impact of prices on analyst expectations. Section 6.3 considers relaxations of these assumptions.

General setting. Let there be I types of days and assume prices impact analyst cash flow expectations heterogeneously across day types, analysts, stocks, and time. Also let the instrument impact prices heterogeneously across day types, analysts, stocks, and time. Then we can decompose the quarterly price change $\Delta p_{a,n,t}$,[53] horizon-h expectation change $\Delta y^h_{a,n,t}$, and instrument $z_{n,t}$ into summations over all day types

$$\begin{aligned} \Delta p_{a,n,t} &= \sum_i \Delta p^i_{a,n,t} \quad \text{and} \quad z_{n,t} = \sum_i z^i_{n,t} \\ \Delta p^i_{a,n,t} &= M^i_{a,n,t} z^i_{n,t} + \epsilon^i_{a,n,t} \\ \Delta y^h_{a,n,t} &= \sum_i \alpha^{i,h}_{a,n,t} \Delta p^i_{a,n,t} + v^h_{a,n,t}. \end{aligned} \tag{17}$$

In this setting, α^{OLS} can be written as

$$\begin{aligned} \alpha_h^{OLS} &= \frac{Cov^{CX}\left(\Delta y^h_{a,n,t}, \Delta p_{a,n,t}\right)}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]} \\ &= \underbrace{\frac{\sum_i Cov^{CX}\left(\alpha^{i,h}_{a,n,t}\Delta p^i_{a,n,t}, \Delta p_{a,n,t}\right)}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}}_{\equiv \alpha_h^{ATE}} + \underbrace{\frac{Cov^{CX}\left(v^h_{a,n,t}, \Delta p_{a,n,t}\right)}{\mathbb{V}^{CX}\left[\Delta p_{a,n,t}\right]}}_{\equiv OVB_h}, \end{aligned} \tag{18}$$

and so the desired covariance proportion is

$$\frac{\sum_i Cov^{CX}\left(\alpha^{i,h}_{a,n,t}\Delta p^i_{a,n,t}, \Delta p_{a,n,t}\right)}{Cov^{CX}\left(\Delta p_{a,n,t}, \Delta y_{a,n,t}\right)} = \frac{\alpha_h^{ATE}}{\alpha_h^{OLS}}, \tag{19}$$

where α_h^{ATE} represents the impact of average price changes on analyst expectations.

Sufficient conditions. As Appendix G.1 proves, $\alpha_h^{2SLS} \cdot b^{meta}$ recovers α_h^{ATE} and $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta}$ recovers desired covariance proportion (19) under the following three conditions (which Section 6.3 relaxes).

Assumption 1 (*Sufficient Conditions for Recovery of Covariance Proportion (19)*)**.** First, there is no estimation noise in α_h^{2SLS}. Second, the 2SLS estimates are uncorrelated with the omitted variable bias term in (18) across all horizons h. Third, the ratio $\theta_h = \alpha_h^{ATE}/\alpha_h^{2SLS}$ is fixed across all horizons h.

Proposition 1 (*Covariance Proportion Recovery in General Setting*)**.** *Under Assumption 1, $\alpha_h^{2SLS} \cdot b^{meta}$ recovers α_h^{ATE} and $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta}$ recovers desired covariance proportion (19) in general setting (17).*

The intuition for Proposition 1 is simple. Following (18), α_h^{OLS} can in general be written as

$$\alpha_h^{OLS} = \alpha_h^{2SLS} \cdot \underbrace{\frac{\alpha_h^{ATE}}{\alpha_h^{2SLS}}}_{\equiv \theta_h} + OVB_h.$$

In general, θ_h is the ratio of the average treatment effect to the local average treatment effect captured by α_h^{2SLS}. So in general, the product $\alpha_h^{2SLS} \cdot \theta_h$ recovers α_h^{ATE}, the numerator of desired covariance proportion (19). Under Assumption 1, b^{meta} recovers the constant $\theta_h = \theta$. Thus, $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta}$ recovers the desired covariance proportion $\alpha_h^{ATE}/\alpha_h^{OLS}$ from (19).

6.3. *Robustness*

This section discusses how relaxations of the three conditions in Assumption 1 impact estimates of the desired covariance proportion (19).

Estimation noise in α_h^{2SLS}. Estimation noise in α_h^{2SLS} acts as measurement error in meta-regression (16) and creates attenuation bias. Thus, b^{meta} underestimates θ, and so $\alpha_h^{2SLS}/\alpha_h^{OLS} \cdot b^{meta}$ underestimates covariance proportion (19).

[53] Note that since $\log(1+x) \approx x$, we have $1 + \Delta p_{a,n,t} = \prod_i \left(1 + \Delta p^i_{a,n,t}\right) \leftrightarrow \Delta p_{a,n,t} \approx \sum_i \Delta p^i_{a,n,t}$.

Pancost and Schaller (2024) provide a simple method to correct for estimation noise in α_h^{2SLS} when estimating θ. Specifically, decompose estimates $\hat{\alpha}_h^{2SLS}$ of α_h^{2SLS} into

$$\hat{\alpha}_h^{2SLS} = \alpha_h^{2SLS} + \zeta_h \epsilon_h,$$

where ϵ_h a mean-zero i.i.d. random variable with unit variance that reflects estimation noise. Then using $\mathbb{V}\left[\hat{\alpha}_h^{2SLS}\right] = \mathbb{V}\left[\alpha_h^{2SLS}\right] + \mathbb{E}\left[\zeta_h^2\right]$, one can remove the attenuation bias from b^{meta}:

$$b^{meta,adj} = b^{meta} \frac{\mathbb{V}\left[\hat{\alpha}_h^{2SLS}\right]}{\mathbb{V}\left[\hat{\alpha}_h^{2SLS}\right] - \mathbb{E}\left[\zeta_h^2\right]}.$$

The desired covariance proportion can then be computed as $\alpha_h^{2SLS} / \alpha_h^{OLS} \cdot b^{meta,adj}$.

Appendix G.2.1 provides the details of this procedure and Table G11 reports the results. After accounting for estimation noise in the 2SLS estimates, the proportions of the covariances of prices with LTG and one to four year EPS expectations explained by the impact of prices on analyst expectations rise from 60% and 40%, respectively, to 92% and 59%. Similarly, as detailed in Table G11, the estimated $\alpha_h^{ATE} = \alpha_h^{2SLS} \cdot b^{meta,adj}$ after accounting for estimation noise in the 2SLS estimates rise from the baseline 2.7 and 11.1 basis points for the LTG and one to four-year EPS expectations, respectively, to 4.0 and 16.7 basis points.

Correlation of α_h^{2SLS} with omitted variable bias. A correlation between α_h^{2SLS} and the omitted variable bias term in α_h^{OLS} (OVB_h in (18)) creates omitted variable bias in the meta-regression (16). Thus, a positive (negative) correlation between these objects leads b^{meta} to overestimate (underestimate) θ, and $\alpha_h^{2SLS} / \alpha_h^{OLS} \cdot b^{meta}$ to overestimate (underestimate) covariance proportion (19).

To address this concern, in Appendix G.2.2 I calculate the values of θ implied by b^{meta} under different assumptions about the correlation of α_h^{2SLS} with OVB_h. I then use these θ values to calculate the desired covariance proportion. As Figure G24 demonstrates, for correlations between α_h^{2SLS} and OVB_h of less than 0.6, the implied covariance proportion point estimates for the LTG and one to four year EPS expectations remain positive. These point estimates are not statistically significantly larger than zero for correlations above 0.2 due to large standard errors arising from uncertainty in the estimated b^{meta} coefficient. Similarly, as Figure G25 demonstrates, the implied α_h^{ATE} point estimates for the LTG and one to four year EPS expectations remain positive for correlations less than 0.6. Due to large standard errors, these point estimates are not statistically significantly larger than zero for correlations above 0.2.

Estimation noise in α_h^{2SLS} and correlation of α_h^{2SLS} with omitted variable bias. Figure G26 in Appendix G.2.3 displays the implied covariance proportions for the LTG and one to four year EPS expectations when accounting for *both* estimation noise in α_h^{2SLS} and potential correlation between α_h^{2SLS} and OVB_h. In this specification, the covariance proportion point estimates remain positive for all correlations up to 0.8. The estimates are not statistically significantly larger than zero for correlations above 0.2 due to large standard errors arising from uncertainty in the estimated b^{meta} coefficient. Similarly, as Figure G.27 demonstrates, the implied α_h^{ATE} point estimates for the LTG and one to four year EPS expectations remain positive for all correlations up to 0.8, while the estimates are not statistically significantly larger than zero for correlations above 0.2.

θ_h varies across horizons. If θ_h varies across horizons h and is uncorrelated with α_h^{2SLS}, but the other two conditions in Assumption 1 are satisfied, then the meta-regression (16) recovers the average θ_h: $b^{meta} = \mathbb{E}[\theta_h]$. Using the average θ_h to calculate the horizon-specific covariance proportions implies some will be overestimated (those for which $\theta_h < \mathbb{E}[\theta_h]$) and some will be underestimated (those for which $\theta_h > \mathbb{E}[\theta_h]$). In particular, between the covariance proportions in Fig. 7 (60% and 40% for the LTG and annual EPS expectations, respectively), one will be overestimated and one will be underestimated. Analogously, between the impacts of average price changes on analyst cash flow expectations ($\alpha_h^{ATE} = \alpha_h^{2SLS} \cdot b^{meta} = 2.7$ and 11.1 basis points for the LTG and annual EPS expectations, respectively), one will be overestimated and one will be underestimated. More precise statements about variation in θ_h across horizons require additional structure, and prove an interesting direction for future work.

7. A model with rational investors that matches subjective beliefs data

This section illustrates how models in which investors have rational expectations can be consistent with stylized facts in subjective beliefs data. Specifically, investors have rational expectations and discount-rate variation creates excess volatility and return predictability. However, investors also have private information that motivates analysts (who seek to forecast cash flows) to learn from prices. Since prices reflect discount rate variation, learning from prices introduces discount rate variation into analyst cash flow expectations. As a result, analyst cash flow expectations predict future returns. By inadvertently attributing discount-rate driven price variation to cash flow news, analysts also form subjective expected returns that correlate weakly with prices. Lastly, analysts overreact to prices, which creates predictable forecast errors. All proofs are in Appendix H.1.

7.1. Model setup

A rational representative investor has private information and sets prices. A representative analyst forecasts cash flows and learns from prices, but has no direct impact on prices.

7.1.1. Setup: Asset pricing block

The representative investor has rational expectations: he observes all shocks and knows all parameters.

Stochastic discount factor. The representative investor's log stochastic discount factor (SDF) is

$$m_{t+1} = -r_f - \frac{1}{2}\gamma^2\sigma^2 - \gamma\epsilon_{t+1},$$

as in Delao et al. (2023). r_f is the log risk-free rate and ϵ_{t+1} is an i.i.d. aggregate shock with variance σ^2.

Cash flows. Log dividend growth for stock n is

$$\begin{aligned} \Delta d_{n,t+1} &= \mu_d + v_{n,t}^{pub} + v_{n,t}^{priv} + \beta_{n,t}\epsilon_{t+1} + u_{n,t+1} \\ v_{n,t+1}^i &= \rho_i v_{n,t}^i + \epsilon_{n,t+1}^i, i \in \{pub, priv\} \\ \beta_{n,t+1} &= 1 + \phi\left(\beta_{n,t} - 1\right) + \epsilon_{n,t+1}^{\beta}. \end{aligned} \tag{20}$$

Dividend growth has two persistent components: a public signal $v_{n,t}^{pub}$ observed by the analyst and investor and a private signal $v_{n,t+1}^{priv}$ observed only by the investor. The private signal motivates the analyst to learn from prices. Note that any private signal the *analyst* has and communicates to the investor acts as a public signal observed by both agents. Dividend growth has a time-varying aggregate-shock exposure ($\beta_{n,t}$) that tractably creates discount rate variation by generating time-varying covariances of cash flows with the SDF. One could generate discount rate variation via other mechanisms as well (e.g. time-varying volatility).

Equilibrium prices. The Campbell and Shiller (1988) approximation yields log price–dividend ratio[54]:

$$\log\left(P_{n,t}/D_{n,t}\right) \approx A_0 + A_1 v_{n,t}^{pub} + A_2 v_{n,t}^{priv} + A_3\beta_{n,t}. \tag{21}$$

[54] As discussed in Appendix H.1.1, the full log price–dividend ratio also depends on $\beta_{n,t}^2$, but I linearize to make the analyst's learning problem tractable. Simulations confirm that, under the estimated parameters in Section 7.2, the approximation error between the full $\log\left(P_{n,t}/D_{n,t}\right)$ and its linearized version is less than 1% (see Appendix H.1.1 for details).

Constants $A_{\_}$ (defined in Appendix H.1.1) are functions of the underlying structural parameters. Prices reflect expected dividend growth (driven by $v^{pub}_{n,t}$ and $v^{priv}_{n,t}$) and discount rates (driven by $\beta_{n,t}$). Discount rate variation creates excess volatility and return predictability.

7.1.2. Setup: Analyst learning block

There is a representative analyst who forecasts cash flows, but does not know the investor's private information $v^{priv}_{n,t}$, the time-varying betas $\beta_{n,t}$, or the realized growth shock $u_{n,t}$. The analyst only observes the public signal $v^{pub}_{n,t}$ and aggregate shock ϵ_t. Thus, he views price and realized dividend growth as a *noisy signals* of $v^{priv}_{n,t}$ and learns from both. He treats all parameters as known; there is no parameter learning.

I first discuss how learning from prices yields analyst cash flow expectations that predict returns, and return expectations that covary weakly with prices. I then discuss how overreaction to prices yields predictable forecast errors. All proofs are in Appendix H.1.2.

Analyst learning from prices. The analyst learns $v^{priv}_{n,t}$ using the Kalman filter. Let

$$s^p_{n,t} \equiv \log(P_{n,t}/D_{n,t}) - A_0 - A_1 v^{pub}_{n,t} = A_2 v^{priv}_{n,t} + A_3 \beta_{n,t} \tag{22}$$
$$s^d_{n,t} \equiv \Delta d_{n,t} - \mu_d - v^{pub}_{n,t-1} = v^{priv}_{n,t-1} + \beta_{n,t-1}\epsilon_t + u_{n,t},$$

be the two signals: the residual log price–dividend ratio ($s^p_{n,t}$) and residual realized dividend growth ($s^d_{n,t}$), both purged of the public signal. Using these signals, the analyst forms an expectation for $v^{priv}_{n,t}$[55]

$$\mathbb{E}^A_t\left[v^{priv}_{n,t}\right] = \rho_{priv}\mathbb{E}^A_{t-1}\left[v^{priv}_{n,t-1}\right] + \alpha\left(s^p_{n,t} - \hat{s}^p_{n,t}\right) + \lambda\left(s^d_{n,t} - \hat{s}^d_{n,t}\right), \tag{23}$$

which informs his total dividend growth expectation

$$\mathbb{E}^A_t\left[\Delta d_{n,t+1}\right] = \mu_d + A_1 v^{pub}_{n,t} + A_2 \mathbb{E}^A_t\left[v^{priv}_{n,t}\right].$$

$\hat{s}^p_{n,t}$ and $\hat{s}^d_{n,t}$ are the predicted signal values given the analyst's expectations of $v^{priv}_{n,t-1}$ and $\beta_{n,t-1}$ from $t-1$. The A superscript indicates the expectation is taken under the analyst's subjective beliefs.

α in (23) is the Kalman gain on price and the model counterpart of the impact of prices on analyst cash flow expectations I measure empirically. Intuitively, α is larger if price is more informative about the investor's private information (i.e. the variance of $v^{priv}_{n,t}$ is larger relative to that of $\beta_{n,t}$). Additionally, α is smaller for higher persistence (ρ_{priv}) of $v^{priv}_{n,t}$, since large price changes correspond to smaller $v^{priv}_{n,t}$ changes.

λ in (23) is the Kalman gain on realized dividend growth.[56] Similar equations (in Appendix H.1.2) describe how the analyst updates expectations of $\beta_{n,t}$.

By learning from prices, the analyst forms cash flow expectations that predict future returns because discount rate variation in the price signal $s^p_{n,t}$ ($\beta_{n,t}$ in (22)) enters his cash flow expectations.

Moreover, by inadvertently attributing discount-rate driven price variation to cash flow news, analysts also form subjective expected returns that correlate weakly with prices. Specifically, price–dividend ratio variation can be decomposed into the proportion explained by subjective expected returns versus the proportion explained by subjective cash flow expectations:

$$1 = -\underbrace{\frac{Cov\left(\log\left(P_{n,t}/D_{n,t}\right), \sum_{s=0}^{\infty}\kappa_1^s\mathbb{E}^A_t\left[r_{n,t+s+1}\right]\right)}{V\left[\log\left(P_{n,t}/D_{n,t}\right)\right]}}_{\text{Proportion Explained by Subjective Expected Returns}} + \underbrace{\frac{Cov\left(\log\left(P_{n,t}/D_{n,t}\right), \sum_{s=0}^{\infty}\kappa_1^s\mathbb{E}^A_t\left[\Delta d_{n,t+s+1}\right]\right)}{V\left[\log\left(P_{n,t}/D_{n,t}\right)\right]}}_{\text{Proportion Explained by Subjective Cash Flow Expectations}}.$$

Learning from prices raises the covariance of prices with subjective cash flow expectations and so suppresses the covariance with subjective expected returns.

Overreaction to prices. To generate predictable forecast errors, I assume the analyst overreacts to prices, which can be microfounded with naive inference (Glaeser and Nathanson, 2017), diagnostic expectations (Bordalo et al., 2021), or partial equilibrium thinking (Bastianello and Fontanier, 2021, 2024). To keep the model as simple as possible, I capture overreaction to prices by modeling the analyst as having biased beliefs about the shock variances of the three unobserved quantities: $v^{priv}_{n,t}$, $\beta_{n,t}$, and $u_{n,t}$. Specifically, the analyst believes $\epsilon^{priv}_{n,t}$, $\epsilon^{\beta}_{n,t}$, and $u_{n,t}$ are i.i.d. with variances:

$$\hat{\sigma}^2_{priv} \equiv \sigma^2_{priv}\cdot(1+\omega)$$
$$\hat{\sigma}^2_{\beta} \equiv \sigma^2_{\beta}\cdot(1-\psi(\omega))$$
$$\hat{\sigma}^2_{u} \equiv \sigma^2_{u} + \pi(\omega).$$

ω controls how much the analyst overestimates the variance of $v^{priv}_{n,t}$. $\psi(\cdot)$ and $\pi(\cdot)$ are functions of ω (defined in Appendix H.1.2) such that the analyst does not misperceive the total variances of $s^p_{n,t}$ and $s^d_{n,t}$. Following the literature, this bias is persistent: analysts do *not* learn the true shock variances over time.

For $\omega = 0$, forecast errors are not predictable since the Kalman filter optimally uses the signals. For $\omega > 0$, however, α is "too large" since the analyst overestimates how informative about $v^{priv}_{n,t}$ the price signal is. Thus, he raises $\mathbb{E}^A_t[v^{priv}_{n,t}]$ "too much" when price rises, and so prices (negatively) predict forecast errors.

This overreaction also suppresses the covariance of prices with subjective expected returns because the analyst underestimates discount rate variation. For sufficiently large ω, $\hat{\sigma}^2_{\beta} = 0$: the analyst believes subjective expected returns are constant and so do not covary with prices.

This overreaction also amplifies how strongly analyst cash flow expectations predict future returns because higher α admits more discount rate variation into analyst cash flow expectations.

7.2. Estimation and model performance

I estimate the model and show that it quantitatively matches several common subjective belief and cross-sectional asset pricing moments. Thus, the impact of prices on analyst cash flow expectations can partially reconcile subjective beliefs data and models featuring investors with rational expectations. Since I identify α in the cross section of stocks, I target other cross-sectional moments for consistency.[57] Table 6 displays the parameters as well as the targeted and model-implied moments. Appendix H.2 provides estimation details.

7.2.1. Targeted moments

I externally calibrate three parameters and estimate the other nine parameters via generalized method of moments to jointly target ten moments.

Externally calibrated parameters. I calibrate $r_f = 0.043$ and $\mu_d = 0.063$ to annual average one-year Treasury bill returns and firm-level earnings growth, and $\sigma_u = 0.23$ to aggregate earnings growth volatility.

[55] All shocks are i.i.d. so there is no learning across stocks (another stock m's price tells the analyst nothing about $v^{priv}_{n,t}$).

[56] In general, the gain parameters can time vary as they depend on aggregate shock ϵ_t. Yet under the estimated parameters in Section 7.2, there is no dependence on ϵ_t as the analyst puts no weight on $s^d_{n,t}$ (i.e. $\lambda = 0$) due to strong overreaction to prices.

[57] When calculating data moments, I remove time fixed effects to focus on the cross section, and stock fixed effects because the model features no permanent differences across stocks.

Asset pricing moments. The first six targeted moments are asset pricing moments.

The first three moments are the variance and first and second autocorrelations of firm-level earnings growth, which depend on the volatilities ($\sigma_{priv}, \sigma_{pub}$, and σ_{β}) and persistences (ρ_{priv}, ρ_{pub}, and ϕ) of the private and public signals as well as $\beta_{n,t}$.[58] The fourth moment is the first autocorrelation of log price–dividend ratio, which depends on all shock volatilities and persistences and is targeted to the first autocorrelation of log market-to-book ratio.[59] The fifth moment is the log market premium, which depends on SDF volatility (controlled by γ and σ) and is targeted to the average CRSP value-weighted index return in excess of the one-year Treasury bill. The sixth moment is cross-sectional return predictability (the proportion of price–dividend ratio variance explained by future returns), which arises only from discount-rate variation and so depends on the volatility and persistence of $\beta_{n,t}$ (σ_{β} and ϕ). I target the 20% of valuation ratio variance explained by future returns documented by Cohen et al. (2003).[60]

Subjective belief moments. The last four targeted moments are subjective belief moments, which all depend on the volatility and persistence of the private signal (σ_{priv} and ρ_{priv}) and $\beta_{n,t}$ (σ_{β} and ϕ), as well as on the overreaction parameter ω.

The first moment is α. Since the model α in (23) represents the impact of average price changes on one-year analyst cash flow growth expectations, it maps to α_1^{ATE}: the average treatment effect of prices on one-year analyst expectations.[61] Thus, I target $\alpha = \alpha_1^{ATE} = 0.12$, calculated as discussed in Section 6.1 (see Appendix Table G10 for details).[62]

The second moment is the predictability of analyst forecast errors: the coefficient in a predictive regression of future one-year forecast errors on current log price–dividend ratio. I target 40% of the scaled empirical covariance between the log market-to-book ratio and one-year EPS forecast errors given the result from Section 6 that 40% of the covariance of analyst one-to-four-year forecast errors with prices comes from the impact of prices on analyst expectations. Section 7.3 discusses extensions of this model that can match the full covariance.

The third moment is the predictability of future returns by analyst cash flow expectations. Following previous work that studies the relationship between analyst cash flow expectations and future returns in the cross section of stocks (e.g. Bordalo et al., 2024; Gormsen and Lazarus, 2023), I focus on the ability of long-term analyst cash flow expectations to predict returns. Specifically, I target the model predictability (i.e. predictive regression coefficient) of four-year returns by current four-year analyst cash flow growth expectations to match the empirical predictability of four-year returns by LTG expectations (similar to Bordalo et al., 2024).

The fourth moment is the scaled covariance of analyst subjective expected returns with prices, which I target to zero to show this covariance can be arbitrarily small.[63]

7.2.2. Model performance

The model matches cross-sectional asset pricing moments well, as displayed in Table 6. Although the model-implied equity premium is slightly large, the model matches firm-level cash flow and valuation ratio dynamics, as well as cross-sectional return predictability.

Crucially, the model also matches subjective belief moments, and so illustrates that models featuring investors with rational expectations can be consistent with subjective beliefs data. The estimated overreaction parameter $\omega = 1.88$ implies the analyst overestimates σ^2_{priv} by a factor of about three, which creates overreaction to prices and predictable forecast errors. Furthermore, the analyst's subjective expected returns are constant and do not covary with prices because the analyst attributes all price signal variation to $v^{priv}_{n,t}$, even the variation that is really due to discount rates. Moreover, since the analyst learns from prices that reflect discount rate variation, analyst cash flow expectations reflect discount rate variation and predict future returns. Overreaction to prices amplifies how strongly these expectations predict returns.

7.3. Extensions

The empirical covariance decomposition results from Section 6 attribute 40% of the covariance of prices with analyst forecast errors to the impact of prices on analyst expectations. Thus, I target the model to match 40% of the empirical covariance of prices with analyst forecast errors. An additional bias in analyst expectations, such as overreaction to fundamental signals, would be required for the model to match the remaining 60% of this covariance. To keep the quantitative model as simple as possible, I do not include a second bias. This section outlines an extended model with overreaction to fundamental signals and Appendix H.3 provides a toy model featuring this mechanism.

Consider an extension of the dividend growth dynamics from (20) that features an additional component $v^{unobs}_{n,t}$ that is unobserved by both the representative investor and analyst:

$$\Delta d_{n,t+1} = \mu_d + v^{pub}_{n,t} + v^{priv}_{n,t} + v^{unobs}_{n,t} + \beta_{n,t}\epsilon_{t+1} + u_{n,t+1}.$$

However, both the investor and analyst observe a noisy signal of this unobserved component: $s^{unobs}_{n,t} = v^{unobs}_{n,t} + \epsilon^{unobs}_{n,t}$. Now assume the analyst overreacts to this signal because he misperceives the variance of $\epsilon^{unobs}_{n,t}$ and believe it is lower than in reality (the same way he overreacts to prices because he underestimates the level of discount rate variation). Thus, the signal $s^{unobs}_{n,t}$ will predict analyst forecast errors. Since the investor learns from $s^{unobs}_{n,t}$ as well, price reflects $s^{unobs}_{n,t}$ and so will also predict analyst forecast errors. Thus, analyst overreaction to fundamental signals raises the covariance of analyst forecast errors with prices.

Importantly, analyst overreaction to fundamental signals contributes to the covariance of analyst forecast errors with prices *even if investors are fully rational* (as the toy model in Appendix H.3 demonstrates). However, this covariance is *larger* if investors also overreact to fundamental signals (as Appendix H.3 demonstrates). There remains an open question of whether a model featuring both investors with rational expectations and an empirically realistic degree of analyst overreaction to both fundamental signals and prices can *quantitatively* match the full covariance of prices with analyst forecast errors. It is possible that only a model with biased investors would be able to do so. I leave these interesting possibilities to future work.

[58] I target moments involving four-year earnings growth instead of annual earnings growth moments to better capture the dynamics of persistent components of earnings growth (which matter more for prices). To address potential issues due to an unbalanced panel, I drop stocks with less than ten years of data when computing data moments. See Appendix H.2 for details.

[59] I use data moments involving the market-to-book ratio because it is defined for all stocks, whereas many stocks do not pay dividends, thereby rendering the price–dividend ratio undefined.

[60] Cohen et al. (2003) look at market equity to book equity ratios instead of price–dividend ratios.

[61] The model α maps less cleanly to the α_1^{2SLS} from either instrument, which represent the impact of price changes driven specifically by those instruments on analyst expectations.

[62] Note that a one percent rise in next-period's expected cash flow (which is what α_1^{ATE} measures) is (approximately) the same as a one percentage point rise in expected growth rate. Assume $\mathbb{E}_t\left[D_{n,t+1}\right]$ rises by δ%, so $\mathbb{E}_t\left[\log\left(D_{n,t+1}\right)\right]$ becomes $\mathbb{E}_t\left[\log\left(D_{n,t+1}\right)+\delta\right]$. Then the new expected growth rate is

$$\begin{aligned}&\mathbb{E}_t\left[\log\left(D_{n,t+1}\right)+\delta-\log\left(D_{n,t}\right)\right]\\&\quad=\mathbb{E}_t\left[\log\left(D_{n,t}\left(1+\Delta d_{n,t+1}\right)\right)+\delta-\log\left(D_{n,t}\right)\right]\\&\quad=\mathbb{E}_t\left[\log\left(1+\Delta d_{n,t+1}\right)\right]+\delta\approx\mathbb{E}_t\left[\Delta d_{n,t+1}\right]+\delta,\end{aligned}$$

where the approximation follows from $\log(1+x)\approx x$.

[63] Delao et al. (2023) find this covariance is negative, but smaller than implied by rational expectations.

Table 6
Model Estimation.

Panel (a): Externally Calibrated Parameters			
Parameter	Value		
r_f	0.043		
σ	0.23		
μ_d	0.063		
Panel (b): Estimated Parameters			
Parameter	Value		
γ	1.73		
σ_{pub}	0.02		
σ_{priv}	0.001		
σ_β	0.03		
σ_u	0.30		
ρ_{pub}	0.48		
ρ_{priv}	0.89		
ϕ	0.91		
ω	1.88		
Panel (c): Target and Model Moments			
	Moment	Target	Model
$\mathbb{V}[G_{n,t}]$	Cross-Sectional Annual EPS Growth Variance	0.36	0.36
$\frac{Cov(G_{n,t+1},G_{n,t})}{\mathbb{V}[G_{n,t}]}$	1st Order Annual EPS Growth Autocorrelation	0.70	0.75
$\frac{Cov(G_{n,t+1},G_{n,t-1})}{\mathbb{V}[G_{n,t}]}$	2nd Order Annual EPS Growth Autocorrelation	0.37	0.50
$\frac{Cov(\log(P_{n,t}/D_{n,t}),\log(P_{n,t-1}/D_{n,t-1}))}{\mathbb{V}[\log(P_{n,t}/D_{n,t})]}$	1st Order Valuation Ratio Autocorrelation	0.64	0.63
$\log\mathbb{E}[r_{m,t}-r_f]$	Log Equity Premium	0.06	0.09
$-\frac{\sum_{s=1}^{15}\kappa_1^s Cov\left(r_{n,t+s},\log(P_{n,t}/D_{n,t})\right)}{\mathbb{V}[\log(P_{n,t}/D_{n,t})]}$	Level of Cross-Sectional Return Predictability	0.20	0.19
$\frac{Cov(\log(P_{n,t}/D_{n,t}),\Delta d_{n,t+1}-\mathbb{E}_t^A[\Delta d_{n,t+1}])}{\mathbb{V}[\log(P_{n,t}/D_{n,t})]}$	Analyst Forecast Error Predictability	−0.01	−0.03
$\frac{Cov(\mathbb{E}_t^A[G_{n,t}],\sum_{s=0}^{4} r_{n,t+1+s})}{\mathbb{V}[\mathbb{E}_t^A[G_{n,t}]]}$	Return Predictability by Analyst Expectations	−0.15	−0.18
$-\frac{\sum_{s=1}^{15}\kappa_1^s Cov\left(\mathbb{E}_t^A[r_{n,t+s}],\log\left(P_{n,t}/D_{n,t}\right)\right)}{\mathbb{V}[\log(P_{n,t}/D_{n,t})]}$	Covariance of Subjective Expected Returns with Prices	0	0
α	Impact of Prices on Analyst Cash Flow Expectations	0.12	0.11

This table reports the externally calibrated parameters in Panel (a), the estimated parameters in Panel (b), and the targeted and model-implied moments in Panel (c). $G_{n,t}$ is four-year dividend growth ($G_{n,t}=\sum_{s=1}^{4}\Delta d_{n,t+s}$). κ_1 is the log-linearization constant from Campbell and Shiller (1988) ($\kappa_1 = 1/\left(1+\exp\left[\mathbb{E}\left[\log\left(D_{n,t}/P_{n,t}\right)\right]\right]\right)$). Appendix H.2 provides details.

8. Conclusion

I propose a potential reconciliation of subjective beliefs data with asset pricing models in which investors have rational expectations and discount rate variation drives prices: I demonstrate that *prices impact analyst cash flow expectations*. Using instruments based on Russell index reconstitution and mutual fund flow-induced trading, I quantify in the cross section of equities how much price increases unrelated to cash-flow news raise analyst cash flow expectations. An exogenous 1% price increase driven by these instruments raises analyst long-term earnings growth expectations by 5 basis points and one to four year EPS expectations and forecast errors by 20 to 40 basis points. This mechanism is economically significant: it explains about half of the covariance of prices with analyst expectations and forecast errors.

To illustrate how this mechanism can partially reconcile subjective beliefs data and models featuring investors with rational expectations, I propose an example model that uses this mechanism to match both subjective belief and cross-sectional asset pricing moments. Investors have private information and discount rate variation drives excess volatility and return predictability. This private information motivates analysts to learn from prices as a signal of future cash flows. However, analysts inadvertently learn from discount rate variation in prices as well. As a result, analysts have biased cash flow expectations that differ from those of investors and forecast future returns, as well as return expectations that weakly correlate with prices.

My results suggest models featuring both investors with rational expectations and discount rate variation need not conflict with subjective beliefs data. While beliefs data can be consistent with investors sharing analysts' biased expectations, they can also be consistent with investors having rational expectations and analysts learning from prices. Moreover, these results raise 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 for the beliefs of a large, price-relevant group of investors? Do large groups of investors also learn from prices as analysts do? Do investors have biased cash flow expectations? Are these biased expectations important drivers of asset prices? Ultimately, since the impact of prices on cash flow expectations arises naturally in models with heterogeneous beliefs, my results suggest heterogeneity may be an important feature of subjective beliefs data and analyst expectations may not align with those of investors. Hence, analyst 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.

Declaration of Generative AI and AI-assisted technologies in the writing process

Generative AI Disclosure: During the preparation of this work the author used ChatGPT, Gemini, and Claude in order to improve language and readability. After using these tools/services, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Aditya Chaudhry reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abarbanell, Jeffery S., 1991. Do analysts' earnings forecasts incorporate information in prior stock price changes? J. Account. Econom. 14 (2), 147–165.

Adam, Klaus, Nagel, Stefan, 2023. Expectations data in asset pricing. In: Handbook of Economic Expectations. Elsevier, pp. 477–506.

Angrist, Joshua, Imbens, Guido, 1995. Identification and estimation of local average treatment effects.

Appel, Ian R., Gormley, Todd A., Keim, Donald B., 2016. Passive investors, not passive owners. J. Financ. Econ. 121 (1), 111–141.

Appel, Ian R., Gormley, Todd A., Keim, Donald B., 2019. Standing on the shoulders of giants: The effect of passive investors on activism. Rev. Financ. Stud. 32 (7), 2720–2774.

Appel, Ian, Gormley, Todd A., Keim, Donald B., 2021. Identification using Russell 1000/2000 index assignments: A discussion of methodologies. Crit. Financ. Rev. Forthcoming.

Bastianello, Francesca, Fontanier, Paul, 2021. Partial equilibrium thinking, extrapolation, and bubbles. Working paper.

Bastianello, Francesca, Fontanier, Paul, 2024. Expectations and learning from prices. Rev. Econ. Stud. rdae059.

Van der Beck, Philippe, 2021. Flow-driven ESG returns. Swiss Finance Institute Research Paper, pp. 21–71.

Van der Beck, Philippe, 2022. On the estimation of demand-based asset pricing models. Swiss Finance Institute Research Paper, pp. 22–67.

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

Ben-David, Itzhak, Li, Jiacui, Rossi, Andrea, Song, Yang, 2022. Ratings-driven demand and systematic price fluctuations. Rev. Financ. Stud. 35 (6), 2790–2838.

Berger, Elizabeth A., 2023. Selection bias in mutual fund fire sales. J. Financ. Quant. Anal. 58 (3), 1039–1077.

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

Bianchi, Francesco, Ludvigson, Sydney C., Ma, Sai, 2023. What hundreds of economic news events say about belief overreaction in the stock market. Available at SSRN.

Bordalo, Pedro, Gennaioli, Nicola, Kwon, Spencer Yongwook, Shleifer, Andrei, 2021. Diagnostic bubbles. J. Financ. Econ. 141 (3), 1060–1077.

Bordalo, Pedro, Gennaioli, Nicola, Porta, Rafael La, Andrei, Shleifer, 2019. Diagnostic expectations and stock returns. J. Financ. 74 (6), 2839–2874.

Bordalo, Pedro, Gennaioli, Nicola, Porta, Rafael La, Andrei, Shleifer, 2024. Belief overreaction and stock market puzzles. J. Polit. Econ. 132 (5), 1450–1484.

Bouchaud, Jean-Philippe, Krueger, Philipp, Landier, Augustin, David, Thesmar, 2019. Sticky expectations and the profitability anomaly. J. Financ. 74 (2), 639–674.

Brandon, Rajna Gibson, Wang, Songtao, 2020. Earnings belief risk and the cross-section of stock returns. Rev. Financ. 24 (5), 1107–1158.

Brown, Philip, Foster, George, Noreen, Eric W., 1987. (Cited often with 1985 working paper date). Security analyst multi-year earnings forecasts and the capital market. J. Account. Econom. 9 (1), 61–87.

Campbell, Y., Shiller, Robert J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Rev. Financ. Stud. 1 (3), 195–228.

Chang, Yen-Cheng, Hong, Harrison, Liskovich, Inessa, 2014. Regression discontinuity and the price effects of stock market indexing. Rev. Financ. Stud. 28 (1), 212–246.

Chaudhry, Aditya, 2024. How much do subjective growth expectations matter for asset prices? Available at SSRN 4209688.

Chen, Long, Da, Zhi, Zhao, Xinlei, 2013. What drives stock price movements? The Rev. Financ. Stud. 26 (4), 841–876.

Chevalier, Judith, Ellison, Glenn, 1997. Risk taking by mutual funds as a response to incentives. J. Political Econ. 105 (6), 1167–1200.

Cohen, Randolph B., Polk, Christopher, Vuolteenaho, Tuomo, 2003. The value spread. J. Financ. 58 (2), 609–641.

Couts, Spencer J., Gonçalves, Andrei, Loudis, Johnathan, 2023. The subjective risk and return expectations of institutional investors. Fisher College of Business Working Paper, (2023-03): 014.

Crane, Alan D., Michenaud, Sébastien., Weston, James P., 2016. The effect of institutional ownership on payout policy: Evidence from index thresholds. Rev. Financ. Stud. 29 (6), 1377–1408.

Dahlquist, Magnus, Ibert, Markus, 2024. Equity return expectations and portfolios: Evidence from large asset managers. Rev. Financ. Stud. 37 (6), 1887–1928.

De Bondt, Werner F.M., Thaler, Richard H., 1990. Do security analysts overreact? The Amer. Econom. Rev. 80 (2), 52–57.

Delao, Ricardo, Han, Xiao, Myers, Sean, 2023. The cross-section of subjective expectations: Understanding prices and anomalies. Available at SSRN.

Delao, Ricardo, Myers, Sean, 2021. Subjective cash flow and discount rate expectations. J. Financ. 76 (3), 1339–1387.

Delao, Ricardo, Myers, Sean, 2024. Which subjective expectations explain asset prices?. Rev. Financ. Stud. 37 (6), 1929–1978.

Derrien, François, Kecskés, Ambrus, Thesmar, David, 2013. Investor horizons and corporate policies. J. Financ. Quant. Anal. 48 (6), 1755–1780.

Dessaint, Olivier, Foucault, Thierry, Frésard, Laurent, Matray, Adrien, 2019. Noisy stock prices and corporate investment. Rev. Financ. Stud. 32 (7), 2625–2672.

Di Maggio, Marco, Franzoni, Francesco, Kogan, Shimon, Xing, Ran, 2023. Avoiding idiosyncratic volatility: Flow sensitivity to individual stock returns. National Bureau of Economic Research.

Dubey, Pradeep, Geanakoplos, John, Shubik, Martin, 1987. The revelation of information in strategic market games: A critique of rational expectations equilibrium. J. Math. Econom. 16 (2), 105–137.

Eckbo, B. Espen, Makaew, Tanakorn, Thorburn, Karin S., 2018. Are stock-financed takeovers opportunistic? J. Financ. Econ. 128 (3), 443–465.

Edmans, Alex, Goldstein, Itay, Jiang, Wei, 2012. The real effects of financial markets: The impact of prices on takeovers. J. Financ. 67 (3), 933–971.

Edmans, Alex, Goldstein, Itay, Jiang, Wei, 2015. Feedback effects, asymmetric trading, and the limits to arbitrage. Am. Econ. Rev. 105 (12), 3766–3797.

Forbes, W.P., Skerratt, L.C.L., 1992. Analysts'forecast revisions and stock price movements. J. Bus. Financ. Account. 19 (4), 555–569.

Frankel, Richard, Lee, Charles M.C., 1998. Accounting valuation, market expectation, and cross-sectional stock returns. J. Account. Econ. 25 (3), 283–319.

Frazzini, Andrea, Lamont, Owen A., 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. J. Financ. Econ. 88 (2), 299–322.

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

Gabaix, Xavier, Koijen, Ralph S.J., 2020. In search of the origins of financial fluctuations: The inelastic markets hypothesis. Available at SSRN 3686935.

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

Giglio, Stefano, Maggiori, Matteo, Stroebel, Johannes, Stephen, Utkus, 2021. Five facts about beliefs and portfolios. Am. Econ. Rev. 111 (5), 1481–1522.

Glaeser, Edward L., Nathanson, Charles G., 2017. An extrapolative model of house price dynamics. J. Financ. Econ. 126 (1), 147–170.

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

Goldsmith-Pinkham, Paul, Sorkin, Isaac, Swift, Henry, 2020. Bartik instruments: What, when, why, and how. Am. Econ. Rev. 110 (8), 2586–2624.

Gormsen, Niels Joachim, Lazarus, Eben, 2023. Duration-driven returns. J. Financ. 78 (3), 1393–1447.

Greenwood, Robin, Shleifer, Andrei, 2014. Expectations of returns and expected returns. Rev. Financ. Stud. 27 (3), 714–746.

Grossman, Sanford J., Stiglitz, Joseph E., 1980. On the impossibility of informationally efficient markets. Am. Econ. Rev. 70 (3), 393–408.

Guay, Wayne, Kothari, S.P., Shu, Susan, 2011. Properties of implied cost of capital using analysts' forecasts. Aust. J. Manag. 36 (2), 125–149.

Heath, Davidson, Macciocchi, Daniele, Michaely, Roni, Ringgenberg, Matthew C., 2022. Do index funds monitor? Rev. Financ. Stud. 35 (1), 91–131.

Hellwig, Martin F., 1980. On the aggregation of information in competitive markets. J. Econom. Theory 22 (3), 477–498.

Hribar, Paul, McInnis, John, 2012. Investor sentiment and analysts' earnings forecast errors. Manag. Sci. 58 (2), 293–307.

Ippolito, Richard A., 1992. Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. J. Law Econ. 35 (1), 45–70.

Jackson, Matthew O., 1991. Equilibrium, price formation, and the value of private information. Rev. Financ. Stud. 4 (1), 1–16.

Jin, Lawrence J., Sui, Pengfei, 2022. Asset pricing with return extrapolation. J. Financ. Econ. 145 (2), 273–295.

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

Koijen, Ralph S.J., Yogo, Motohiro, 2019. A demand system approach to asset pricing. J. Political Econ. 127 (4), 1475–1515.

Kyle, Albert S., 1989. Informed speculation with imperfect competition. Rev. Econ. Stud. 56 (3), 317–355.

Landier, Augustin, Thesmar, David, 2020. Earnings expectations during the COVID-19 crisis. Rev. Asset Pricing Stud. 10 (4), 598–617.

Lee, David S., McCrary, Justin, Moreira, Marcelo J., Porter, Jack, 2022. Valid t-ratio inference for IV. Am. Econ. Rev. 112 (10), 3260–3290.

Lee, Charles M.C., Myers, James, Swaminathan, Bhaskaran, 1999. What is the intrinsic value of the dow? J. Financ. 54 (5), 1693–1741.
Lee, Charles M.C., So, Eric C., 2017. Uncovering expected returns: Information in analyst coverage proxies. J. Financ. Econ. 124 (2), 331–348.
Lee, Charles M.C., Swaminathan, Bhaskaran, 2000. Price momentum and trading volume. J. Financ. 55 (5), 2017–2069.
Li, Jiacui, 2022. What drives the size and value factors? Rev. Asset Pricing Stud. 12 (4), 845–885.
Li, Jian, Fu, Zhiyu, Chaudhary, Manav, 2022. Corporate bond elasticities: Substitutes matter. Available at SSRN 4292266.
Lou, Dong, 2012. A flow-based explanation for return predictability. Rev. Financ. Stud. 25 (12), 3457–3489.
Lou, Xiaoxia, Wang, Albert Y., 2018. Flow-induced trading pressure and corporate investment. J. Financ. Quant. Anal. 53 (1), 171–201.
Lys, Thomas, Sohn, Sungkyu, 1990. The association between revisions of financial analysts' earnings forecasts and security-price changes. J. Account. Econ. 13 (4), 341–363.
Malkiel, Burton G., 1970. The valuation of public utility equities. Bell J. Econ. Manag. Sci. 1 (1), 143–160.
McCarthy, Odhrain, Hillenbrand, Sebastian, 2021. Heterogeneous beliefs and stock market fluctuations. Available at SSRN 3944887.
Mendel, Brock, Shleifer, Andrei, 2012. Chasing noise. J. Financ. Econ. 104 (2), 303–320.
Miller, Jeffrey S., Sedor, Lisa M., 2014. Do stock prices influence analysts' earnings forecasts? Behav. Res. Account. 26 (1), 85–108.
Nagel, Stefan, Xu, Zhengyang, 2021. Asset pricing with fading memory. Rev. Financ. Stud. 34 (6), 2838–2873.
Nagel, Stefan, Xu, Zhengyang, 2023. Dynamics of subjective risk premia. J. Financ. Econ. 150 (2), 103713.
Norli, Øyvind, Ostergaard, Charlotte, Schindele, Ibolya, 2015. Liquidity and shareholder activism. Rev. Financ. Stud. 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, Schaller, Garrett, 2024. Measuring measurement error. Available at SSRN 4045772.
Pavlova, Anna, Sikorskaya, Taisiya, 2023. Benchmarking intensity. Rev. Financ. Stud. 36 (3), 859–903.
Phillips, Gordon M., Zhdanov, Alexei, 2013. R & D and the incentives from merger and acquisition activity. Rev. Financ. Stud. 26 (1), 34–78.
Rafael, La Porta, 1996. Expectations and the cross-section of stock returns. J. Financ. 51 (5), 1715–1742.
Schmidt, Cornelius, Fahlenbrach, Rüdiger, 2017. Do exogenous changes in passive institutional ownership affect corporate governance and firm value? J. Financ. Econ. 124 (2), 285–306.
Securities, U.S. and Exchange Commission, 2023a. Form N-CSR.
Securities, U.S. and Exchange Commission, 2023b. 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? Audit. Financ. 20 (2), 147–166.
Sirri, Erik R., Tufano, Peter, 1998. Costly search and mutual fund flows. J. Financ. 53 (5), 1589–1622.
Subrahmanyam, Avanidhar, Titman, Sheridan, 2001. Feedback from stock prices to cash flows. J. Financ. 56 (6), 2389–2413.
Sulaeman, Johan, Wei, Kelsey D., 2019. Sell-side analysts and stock mispricing: evidence from mutual fund flow-driven trading pressure. Manag. Sci. 65 (11), 5427–5448.
Wardlaw, Malcolm, 2020. Measuring mutual fund flow pressure as shock to stock returns. J. Financ. 75 (6), 3221–3243.
Wei, Wei, Young, Alex, 2021. Selection bias or treatment effect? A re-examination of Russell 1000/2000 index reconstitution. Crit. Financ. Rev. 10.
Wharton Research Data Services, 2008. TF Estimates Glossary. (Accessed on 26 March 2023).
Wu, Liuren, 2018. Estimating risk-return relations with analysts' price targets. J. Bank. Financ. 93, 183–197.
Zhu, Qifei, 2020. The missing new funds. Manag. Sci. 66 (3), 1193–1204.
Zuo, Luo, 2016. The informational feedback effect of stock prices on management forecasts. J. Account. Econ. 61 (2–3), 391–413.