On the Recovery and Usage of Demand Elasticities in Dynamic Settings*

November 18, 2025

Abstract

We study how to estimate and interpret demand elasticities in dynamic settings where prices and flows are jointly determined. Because price changes necessarily alter expected returns or cash flows, there is no single context-free elasticity: measured elasticities depend on which expectations adjust. We formalize static, immediate, and dynamic elasticities and show that the dynamic elasticity equals the inverse of the equilibrium price multiplier. A tractable linear model delivers testable comparative statics linking price multipliers to risk, persistence, systematic exposure, and surprise. Empirically, we confirm these patterns and show that structural demand can be recovered from reduced-form price responses even under persistent shocks.

KEYWORDS: demand elasticity, price impact, demand system asset pricing

JEL CLASSIFICATION: G11, G12, G14.

Demand system asset pricing (DSAP) has become a central framework for studying how investor flows shape asset prices and for conducting counterfactual analysis. By modeling investor demand as a function of prices, DSAP provides a flexible empirical structure for linking portfolio behavior to market equilibrium. This approach has been applied to settings such as ESG (Koijen, Richmond, Yogo, et al., 2019; Van der Beck, 2021), passive intermediation (Haddad, Huebner, and Loualiche, 2025b), household participation (Davis, Knüpfer, Soerlie Kvaerner, Sen Dogan, and Vokata, 2024), and anomaly returns (Tamoni, Sokolinski, and Li, 2024), among others. The growing influence of this approach reflects its promise: once estimated, demand elasticities can be used to infer price impact of flows and enable counterfactual analyses that quantify how changes in investor composition or regulation would reshape asset prices.

As the framework has diffused across contexts, two fundamental questions have become increasingly important. First, can demand elasticities be reliably estimated when prices and flows are dynamic—when instruments for exogenous demand shocks are persistent and propagate through expectations of future prices and returns? Second, when elasticities are used for counterfactual analysis, how closely do they correspond to the structural objects that govern equilibrium price responses? These questions are critical because in finance, price today is not a static object: any price movement affects, and is affected by, expectations of future dividends and discount rates.

Our paper addresses both issues. We show that dynamics create an inherent tension between the elasticity we can estimate and the one relevant for counterfactual analysis. In particular, price changes induced by persistent instruments propagate along the entire path of expected returns, so the resulting elasticity reflects both immediate and intertemporal adjustments rather than a one-period response. We formalize this mapping by distinguishing static, immediate, and dynamic elasticities, each corresponding to a different assumption about which fundamentals adjust when prices move.

Among these, the dynamic elasticity governs the equilibrium price multiplier in the presence of persistent demand shocks.

Because prices are forward-looking, a movement today is generally associated with the entire path of expected future returns and dividends, not just current valuations. Consequently, there is no single notion of “the” price elasticity in finance: the object recovered from data depends on which expectations are allowed to adjust. Our first main result formalizes this dependence and shows that a context-invariant elasticity does not exist. Different sources of price variation—cash-flow news, short-horizon discount-rate news, or long-horizon revaluations—induce different objects that are all consistent with the same structural model.

The problem of intertemporal linkages becomes especially acute when the exogenous demand that serves as the instrument for price changes is itself persistent. Persistent flows affect not only current prices but also expectations of future prices and returns, blurring the distinction between immediate and long-run adjustments. Empirical elasticities estimated from such instruments therefore embed both contemporaneous and intertemporal responses and cannot be interpreted without an explicit dynamic structure. This concern is especially relevant in practice, as many DSAP applications rely on persistent variation—such as the rise of passive funds or the accumulation of household savings—precisely where static interpretations are least appropriate.

We address these challenges in two steps. First, we provide a general decomposition that clarifies what any empirical elasticity measures once the intertemporal structure of prices is taken into account. We define three coherent elasticity concepts, each determined by which fundamentals are allowed to adjust when prices move. The static return elasticity isolates cash-flow news while holding all discount rates fixed. The immediate return elasticity captures the response to a one-period change in expected returns, holding longer-horizon expectations constant. The dynamic return elasticity allows prices to revalue the entire term structure of expected returns while holding cash flows fixed. These elasticities are not competing definitions but different conditional responses implied by the same underlying demand system. Recognizing this taxonomy reconciles seemingly conflicting interpretations in the literature and clarifies which elasticity governs equilibrium price multipliers.

Second, we develop a tractable linear equilibrium model to characterize how these elasticities are related and what determines their magnitudes. A representative arbitrageur absorbs outside flows with perceived return covariance matrix Σ, while the outside-investment process follows an AR(1) with persistence ρ. Solving the pricing recursion yields closed-form expressions linking price multipliers and elasticities. In equilibrium, the dynamic elasticity is exactly the inverse of the price multiplier, whereas the immediate elasticity—which holds longer-horizon price expectations fixed—captures only a partial-equilibrium response and therefore does not correspond to the inverse of the multiplier matrix. The model further shows that both elasticities depend systematically on observable features of the environment. Price multipliers—and therefore inverse elasticities—are larger when risk is higher (through Σ), when flows are more persistent (higher ρ), when flows load on more systematic rather than idiosyncratic components of risk, and when flows are surprises rather than anticipated.

These insights have direct implications for both estimation and counterfactual analysis. For estimation, they show that dynamic settings do not invalidate demand measurement; rather, they define precisely the object that is measured: the dynamic elasticity mapping persistent flow shocks to equilibrium prices. For counterfactuals, they emphasize that applying an estimated elasticity or multiplier to a new setting implicitly assumes that the counterfactual environment shares the same variance, persistence, and systematic structure as the original identification source. When those features differ, price responses will generally differ as well, even if the estimated elasticity remains unchanged.

Empirically, we test the model's comparative statics using non–cash-flow demand shocks and find strong support for each predicted pattern. Flows that are more persistent, more systematic, and more surprising generate larger price multipliers, consistent with the theory. Finally, we show that even when instruments are persistent, the underlying structure of demand can be recovered from reduced-form price responses once the intertemporal nature of price adjustment is taken into account.

We then formalize how reduced-form price dynamics identify the structural demand objects governing portfolio choice. A persistent instrument induces a sequence of observable impulse responses linking outside-investment shocks to future prices. These impulse responses embed intertemporal substitution—how current and future expected prices co-move with the instrument—because equilibrium ties today's price to the expected path of future prices. We show that, under mild stability and invertibility conditions, these reduced-form responses together with the stochastic dynamics of the instrument pin down the structural demand matrices mapping expected prices into holdings. In the benchmark setting, this inversion collapses to the result that the dynamic elasticity equals the inverse price multiplier; in general, it recovers both the contemporaneous and forward-looking components of demand that the multiplier alone does not reveal.

Our contribution is fourfold. First, we clarify why there is no single demand elasticity in finance and provide a taxonomy that organizes empirical estimates by the fundamentals that adjust when prices move. Second, we derive a dynamic equilibrium linking flows, prices, and elasticities, identify the dynamic elasticity as the equilibrium object governing persistent shocks, and characterize the forces shaping its magnitude. Third, we empirically validate these comparative statics and offer guidance for interpreting and applying elasticities in dynamic settings. Fourth, we show how structural demand can be recovered even when identification relies solely on persistent price movements that revalue the entire term structure of expected returns. Together, these results demonstrate that demand elasticities are not ill-defined; they are economic objects whose interpretation depends on the source and persistence of the shocks that move prices.

1 Different Types of Price Elasticities

The object of interest is the price elasticity of portfolio demand. Let p_t denote the $N \times 1$ vector of log prices at time t, d_t the vector of log dividends, r_t the vector of log returns, and q_t the vector of log demand (log number of shares demanded). For $j \geq 1$, let $\mathbb{E}_t x_{t+j}$ denote possibly subjective conditional expectations. The elasticity we seek is the Jacobian

$$-\frac{\partial q_t}{\partial p_t^\top},$$

which answers the question: if current prices move by one percent, by how much does contemporaneous demand respond, along both own- and cross-price directions. In dynamic asset markets, however, price movements cannot be evaluated in isolation. A price change today typically shifts expectations of future returns and cash flows, and these shifts feed back into demand. Any meaningful definition of elasticity must therefore state what is held fixed when prices move.

1.1 An intentionally overdetermined demand

To make the identification problem transparent, suppose we write demand as a function of current prices and the entire forward path of expectations,

$$q_t = q\Big(p_t, \mathbb{E}_t p_{t+1}, \mathbb{E}_t d_{t+1}, \mathbb{E}_t r_{t+1}, \mathbb{E}_t p_{t+2}, \mathbb{E}_t d_{t+2}, \mathbb{E}_t r_{t+2}, \ \ldots\Big), \tag{1}$$

abstracting from changes in higher moments so that only first moments vary over time. The naive elasticity $-\partial q_t / \partial p_t^\top$ computed from (1) implicitly holds fixed the entire sequence $\left(\mathbb{E}_t p_{t+j}, \mathbb{E}_t d_{t+j}, \mathbb{E}_t r_{t+j}\right)_{j \geq 1}$. This thought experiment is internally inconsistent. Prices, expected returns, and expected dividends are tied by accounting identities, so it is incoherent to move p_t while keeping all forward-looking objects unchanged. Even setting aside consistency, most standard preferences imply that investors care about prices only through their implications for expected returns, so price changes that leave expected returns unchanged are typically irrelevant for demand. The conclusion is that the elasticity cannot be defined until we specify which forward-looking objects adjust with prices and which are held fixed.

1.2 Campbell-Shiller approximation and price pass-through

The Campbell-Shiller relation provides the key accounting that links p_t, future dividends, and expected returns. In its convenient log-linear form,

$$p_t - d_t = \kappa + \rho \left(p_{t+1} - d_{t+1}\right) + \Delta d_{t+1} - r_{t+1},$$

for constants κ and ρ from the first-order approximation. Iterating forward under the usual transversality condition yields the long-run representation:

$$p_t = \frac{\kappa}{1-\rho} + (1-\rho) \sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t r_{t+j}. \tag{2}$$

Differentiating (2) with respect to p_t shows that any price movement must pass through expected dividends and/or expected returns:

$$I = (1-\rho)\sum_{j=1}^{\infty}\rho^{j-1}\underbrace{\frac{\partial \mathbb{E}_t d_{t+j}}{\partial p_t^\top}}_{\substack{\text{pass-through to}\\ \text{dividends}}} + \sum_{j=1}^{\infty}\rho^{j-1}\underbrace{\left(-\frac{\partial \mathbb{E}_t r_{t+j}}{\partial p_t^\top}\right)}_{\substack{\text{pass-through to}\\ \text{expected returns}}} \tag{3}$$

where I is the $N \times N$ identity. Thus price changes cannot occur in a vacuum; they necessarily manifest as adjustments in expected cash flows, expected returns, or both. The equation above shows both prices passing through into dividend expectations and expected returns. The sum of these pass-through terms, weighted by the relevant ρ terms, must sum to one.

We show a simple calibration. We use an annual investment horizon and $\rho = 0.9638$ from Cochrane (2008). The diagonal pass-through terms implied by (3) are reported in Table 1. When a price change at t is absorbed exclusively through expected dividends at horizon j, the diagonal mapping equals $\left[(1-\rho)\rho^{j-1}\right]^{-1}$; when it is absorbed exclusively through expected returns at horizon j, it equals $\rho^{-(j-1)}$. Over $j = 1, \ldots, 5$, the table shows how these magnitudes increase with horizon, reflecting the geometric weighting in (3).

Equation (3) and the table make clear that choosing what to hold fixed when prices move selects a specific pass-through path across horizons and across the dividend versus expected-return channels. Elasticities differ precisely because these pass-through patterns differ.

1.3 Example 1: $(p_0 \to r_1, d)$ demand

Consider the following coherent representation of demand that isolates a one-period discount-rate movement when prices change and holds fixed all dividend expectations and all expected returns except at the one-step-ahead horizon. The notation means that p_t changes affect only $\mathbb{E}_t r_{t+1}$,

j	$\dfrac{I}{(1-\rho)\rho^{\,j-1}}$ =	$\dfrac{\partial \mathbb{E}_t d_{t+j}}{\partial p_t^\top}$	$\dfrac{I}{\rho^{\,j-1}}$ =	$\left(-\dfrac{\partial \mathbb{E}_t r_{t+j}}{\partial p_t^\top}\right)$
1	$\dfrac{1}{(1-\rho)\rho^0}$	27.6243	ρ^{-0}	1.0000
2	$\dfrac{1}{(1-\rho)\rho^1}$	28.6619	ρ^{-1}	1.0376
3	$\dfrac{1}{(1-\rho)\rho^2}$	29.7384	ρ^{-2}	1.0765
4	$\dfrac{1}{(1-\rho)\rho^3}$	30.8554	ρ^{-3}	1.1170
5	$\dfrac{1}{(1-\rho)\rho^4}$	32.0143	ρ^{-4}	1.1589

Table 1. Diagonal pass-through magnitudes implied by (3) under $\rho = 0.9638$. The left pair of columns reports the mapping when price changes transmit solely through expected dividends at horizon j; the right pair reports the mapping when they transmit solely through expected returns at horizon j.

and all other expected returns are held fixed. Furthermore, the d in the superscript means the function includes all expected dividends terms, meaning that the partial derivative with respect to prices holds these arguments fixed. Write

$$q_t^{(p_0\to r_1,d)} = q^{(p_0\to r_1,d)}\big(p_t, \mathbb{E}_t d_{t+1}, \mathbb{E}_t d_{t+2}, \mathbb{E}_t r_{t+2}, \mathbb{E}_t d_{t+3}, \mathbb{E}_t r_{t+3}, \ldots\big),$$

with the omitted arguments $\big(\mathbb{E}_t r_{t+1}, \mathbb{E}_t p_{t+1}, \mathbb{E}_t p_{t+2}, \ldots\big)$ adjusting so that (3) holds and only $\mathbb{E}_t r_{t+1}$ moves on the return side when prices move. The corresponding elasticity is

$$-\frac{\partial q_t^{(p_0\to r_1,d)}}{\partial p_t^\top}. \tag{4}$$

In words, (4) measures the response of demand to a price change that maps one-for-one into next-period expected returns, with all dividend expectations and longer-horizon expected returns held fixed. Stated more simply, this is the sensitivity of the investor to a 1% rise in next-period expected returns that occurs simultaneously with a 1% drop in prices.

This elasticity is the same as the elasticity from this demand function:

$$q_t^{(p,d)} = q^{(p,d)}(p_t, \mathbb{E}_t p_{t+1}, \mathbb{E}_t d_{t+1}, \mathbb{E}_t p_{t+2}, \mathbb{E}_t d_{t+2}, ...),$$

with omitted arguments $(\mathbb{E}_t r_{t+1}, \mathbb{E}_t r_{t+2}, ...)$. By Campbell-Shiller, it's easy to see that if future expectations of prices and dividends are held fixed, then a 1% drop in prices corresponds to a 1% rise in next period expected returns, $\mathbb{E}_t r_{t+1}$.

1.4 Example 2: $(p_0 \to r_2, d)$ demand

To capture sensitivity to a purely two-period-ahead discount-rate movement, hold fixed all dividend expectations and all expected returns except at horizon $t+2$:

$$q_t^{(p_0\to r_2,d)} = q^{(p_0\to r_2,d)}\big(p_t, \mathbb{E}_t d_{t+1}, \mathbb{E}_t r_{t+1}, \mathbb{E}_t d_{t+2}, \mathbb{E}_t d_{t+3}, \mathbb{E}_t r_{t+3}, \dots\big),$$

with the omitted expectations $\big(\mathbb{E}_t r_{t+2}, \mathbb{E}_t p_{t+1}, \mathbb{E}_t p_{t+2}, \dots\big)$ adjusting so that (3) holds and only $\mathbb{E}_t r_{t+2}$ moves on the return side. The elasticity

$$-\frac{\partial q_t^{(p_0\to r_2,d)}}{\partial p_t^\top}$$

is the response to a price change that implies a $1/\rho$-scaled movement in expected returns two periods ahead. In other words, if $\rho \approx 0.96$, then this elasticity corresponds to how much, in percentage terms, the investor increases demand when simultaneously prices drop by 1% and expected returns rise two periods from t by 1.04%. If demand is myopic (for example, mean-variance demand in one-period problems), this elasticity is zero by construction since the investor does not react to expected returns beyond the next period.

1.5 Example 3: $(p_0 \to d_1, r)$ demand

To isolate the effect of cash-flow news, hold fixed the entire term structure of expected returns and allow only next-period dividend expectations, $\mathbb{E}_t d_{t+1}$, to adjust with prices:

$$q_t^{(p_0 \to d_1, r)} = q^{(p_0 \to d_1, r)}\big(p_t, \mathbb{E}_t r_{t+1}, \mathbb{E}_t d_{t+2}, \mathbb{E}_t r_{t+2}, \mathbb{E}_t d_{t+3}, \mathbb{E}_t r_{t+3}, \dots\big),$$

where $\big(\mathbb{E}_t d_{t+1}, \mathbb{E}_t p_{t+1}, \mathbb{E}_t p_{t+2}, \dots\big)$ adjust to satisfy (3). The corresponding elasticity

$$-\frac{\partial q_t^{(p_0 \to d_1, r)}}{\partial p_t^\top}$$

measures the change in demand resulting from a price movement that translates into news about next-period dividends while leaving all expected returns unchanged. Quantitatively, using $\rho = 0.9638$, a 1% price change corresponds to a $1\% \times \frac{1}{1-\rho} \approx 27.6\%$ movement in expected dividends at horizon $t + 1$. In words, this elasticity measures the demand response in percentage terms due to a simultaneous 1% drop in prices and 27.6% drop in next-period dividends, holding all else constant. Under standard preferences in which dividends matter only through their implications for returns, this elasticity is zero by construction.

1.6 There is no "one" price elasticity in finance

These examples make clear that there is no single, universal notion of a "price elasticity" in finance. The appropriate elasticity depends entirely on what is held fixed when prices move—that is, on the pass-through structure between prices, expected returns, and expected dividends implied by equilibrium accounting. The $(p_0 \rightarrow r_1, d)$ elasticity captures the response to a one-period discount-rate movement, while the $(p_0 \rightarrow r_2, d)$ and $(p_0 \rightarrow d_1, r)$ elasticities capture distinct forms of the discount-rate two periods away and next-period cash-flow news, respectively. Each is well-defined within its own ceteris paribus clause, but they are fundamentally different objects that will generally not coincide.

Binsbergen, David, and Opp (2025) provide a theoretical definition of elasticity that corresponds precisely to the $(p_0 \rightarrow r_1, d)$ case. They consider the $q_t^{(p,d)}$ representation of demand, which holds fixed expectations of all future prices and dividends. As discussed above, this is equivalent to allowing only next-period expected returns to adjust.

While Binsbergen et al. (2025) clearly articulate their elasticity concept, much of the demand system asset pricing literature has not been as explicit about what is held fixed when prices change. In empirical implementations of demand system asset pricing, authors often speak of "the" elasticity without specifying whether their price movements reflect cash-flow news, short-term discount-rate changes, or long-term revaluations. Because the equilibrium mapping between prices, expected returns, and dividends is mechanical, such ambiguity risks conflating distinct economic objects. A consistent empirical and theoretical dialogue in this literature requires that authors identify precisely which elasticity their design recovers and under what ceteris paribus conditions.

1.7 Indirect demand

A useful perspective is that demand is indirect. In consumer theory, goods deliver utility through their attributes (e.g., Lancaster, 1966)[1] and households combine time and market goods to

[1] Lancaster (1966) writes: "[t]he good, per se, does not give utility to the consumer; it possesses characteristics, and these characteristics give rise to utility."

produce the commodities they value (Becker, 1965).[2] Asset demand is analogous. Investors value characteristics of assets—expected returns and the cash-flow dynamics that shape consumption paths—rather than the assets per se. Prices matter because, through identities such as Campbell-Shiller, they reconfigure these characteristics across horizons. Once this pass-through is made explicit, the many elasticities considered above become natural: each corresponds to a distinct ceteris paribus clause about how prices today map into the characteristics investors actually care about.

2 Instrumented demand elasticities

This section formalizes what empirical designs recover when prices are moved by observable instruments. The key point is that price changes pass through to expected returns and/or expected dividends, so the measured object depends on which characteristics the instrument affects.

2.1 Set-up and definition

Let $z_{t-s} \in \mathbb{R}^N$ be an instrument realized $s \geq 0$ periods before t. The instrumented demand elasticity at horizon (t,s) is defined as

$$\eta_{t,s} = - \underbrace{\frac{dq_t}{dz_{t-s}^{\top}}}_{\text{reduced form}} \left(\underbrace{\frac{dp_t}{dz_{t-s}^{\top}}}_{\text{first stage}} \right)^{-1} . \tag{5}$$

[2] Becker (1965) states: "[h]ouseholds will be assumed to combine time and market goods to produce more basic commodities that directly enter their utility functions. One such commodity is the seeing of a play, which depends on the input of actors, script, theatre and the playgoer's time; another is sleeping, which depends on the input of a bed, house (pills?) and time."

where $dq_t/dz_{t-s}^\top$ is the reduced-form response of demand to the instrument and $dp_t/dz_{t-s}^\top$ is the first-stage response of prices.

To connect (5) to primitives, consider a generic representation

$$q_t = q\big(p_t, x_{0,t}, x_{1,t}, x_{2,t}, \ldots\big), \qquad x_{j,t} \in \{\mathbb{E}_t r_{t+k}, \mathbb{E}_t d_{t+k}, \mathbb{E}_t p_{t+k}, \ldots\}_{k\geq 1}.$$

By the chain rule,

$$\frac{dq_t}{dz_{t-s}^\top} = \frac{\partial q_t}{\partial p_t^\top}\frac{dp_t}{dz_{t-s}^\top} + \sum_{j\geq 0}\frac{\partial q_t}{\partial x_{j,t}^\top}\frac{\partial x_{j,t}}{\partial z_{t-s}^\top}. \tag{6}$$

If $x_{j,t}$ moves only through prices, so that $\partial x_{j,t}/\partial z_{t-s}^\top = (\partial x_{j,t}/\partial p_t^\top)(dp_t/dz_{t-s}^\top)$, substituting (6) into (5) yields

$$\eta_{t,s} = \underbrace{-\frac{\partial q_t}{\partial p_t^\top}}_{\substack{\text{pure demand} \\ \text{elasticity}}} \underbrace{-\sum_{j\geq 0}\frac{\partial q_t}{\partial x_{j,t}^\top}\frac{\partial x_{j,t}}{\partial p_t^\top}}_{\substack{\text{instrumented wedge:} \\ \text{pass-through via} \\ \text{characteristics}}}. \tag{7}$$

Expression (7) shows that an instrumented elasticity generally mixes a direct price sensitivity with a wedge induced by how prices reconfigure characteristics $x_{j,t}$. Different instruments generate different wedges because they imply different pass-through patterns.

Below we consider the following representation of demand

$$q_t = q_t^{(d,r)}\left(\mathbb{E}_t d_{t+1}, \mathbb{E}_t r_{t+1}, \mathbb{E}_t d_{t+2}, \mathbb{E}_t r_{t+2}, \ \ldots\right), \tag{8}$$

which omits p_t because, under the no-bubble condition and the Campbell-Shiller identity, current prices are fully determined by the term structure of expected dividends and expected returns. Including p_t alongside the entire path of $(\mathbb{E}_t d_{t+j}, \mathbb{E}_t r_{t+j})$ would overdetermine the state.

2.2 Static instrumented elasticity (SIE)

A static return, or cash-flow, instrument moves expected dividends while leaving expected returns fixed at all horizons:

$$\frac{\partial \mathbb{E}_t[r_{t+j}]}{\partial z_{t-s}^{\top}} = 0 \quad \text{for all } j \geq 1, \qquad \frac{\partial \mathbb{E}_t[d_{t+j}]}{\partial z_{t-s}^{\top}} \neq 0 \text{ for some } j \geq 1.$$

Using (8),

$$\frac{dq_t^{(d,r)}}{dz_{t-s}^{\top}} = \sum_{j\geq 1} \frac{\partial q_t^{(d,r)}}{\partial \mathbb{E}_t[d_{t+j}]^{\top}} \frac{\partial \mathbb{E}_t[d_{t+j}]}{\partial z_{t-s}^{\top}}. \tag{9}$$

Substituting (9) into (5) gives

$$\eta_{t,s}^{\text{SIE}} = -\left(\sum_{j\geq 1} \frac{\partial q_t^{(d,r)}}{\partial \mathbb{E}_t[d_{t+j}]^{\top}} \frac{\partial \mathbb{E}_t[d_{t+j}]}{\partial z_{t-s}^{\top}}\right)\left(\frac{dp_t}{dz_{t-s}^{\top}}\right)^{-1}.$$

Hence SIE measures the elasticity of demand with respect to cash-flow news, holding discount rates fixed. Under mean-variance demand, or any specification in which $q_t^{(d,r)}$ depends only on expected returns and not directly on expected dividends, $\partial q_t^{(d,r)}/\partial \mathbb{E}_t[d_{t+j}]^{\top} = 0$ for all j, implying $\eta_{t,s}^{\text{SIE}} = 0$. This elasticity is zero in these settings because investors with these preferences do not care about cash flows per se, but returns. Since the instrument does not affect expected returns, the investor is not sensitive to these price changes at all.

2.3 Immediate instrumented elasticity (IIE)

An immediate return instrument generates a pure contemporaneous discount-rate movement: it changes p_t and only next-period expected returns, while leaving expected dividends and all longer-horizon expected returns fixed:

$$\frac{\partial \mathbb{E}_t[d_{t+j}]}{\partial z_{t-s}^\top} = 0 \text{ for all } j \geq 1, \qquad \frac{\partial \mathbb{E}_t[r_{t+j}]}{\partial z_{t-s}^\top} = 0 \text{ for } j \geq 2.$$

By Campbell-Shiller, $-\partial \mathbb{E}_t[r_{t+1}]/\partial p_t^\top = I_N$. The reduced-form response under (8) is

$$\frac{dq_t^{(d,r)}}{dz_{t-s}^\top} = \frac{\partial q_t^{(d,r)}}{\partial \mathbb{E}_t[r_{t+1}]^\top} \frac{\partial \mathbb{E}_t[r_{t+1}]}{\partial z_{t-s}^\top}.$$

Using $\partial \mathbb{E}_t[r_{t+1}]/\partial z_{t-s}^\top = (\partial \mathbb{E}_t[r_{t+1}]/\partial p_t^\top)(dp_t/dz_{t-s}^\top) = -I_N\,(dp_t/dz_{t-s}^\top)$ in (5) yields the simple form

$$\eta_{t,s}^{\text{IIE}} = \frac{\partial q_t^{(d,r)}}{\partial \mathbb{E}_t[r_{t+1}]^\top}.$$

Equivalently, in the (p,d) representation that includes prices and holds future dividends and prices fixed, $\eta_{t,s}^{\text{IIE}} = -\partial q_t^{(p,d)}/\partial p_t^\top$. Thus IIE coincides with the marginal demand for a one-period expected return, the notion emphasized by designs that isolate short-lived pricing deviations.

2.4 Dynamic instrumented elasticity (DIE)

A dynamic instrument shifts the term structure of expected returns while holding expected dividends fixed:

$$\frac{\partial \mathbb{E}_t[d_{t+j}]}{\partial z_{t-s}^\top} = 0 \text{ for all } j \geq 1, \qquad \frac{\partial \mathbb{E}_t[r_{t+j}]}{\partial z_{t-s}^\top} \neq 0 \text{ for some } j \geq 1.$$

The reduced-form response under (8) is

$$\frac{dq_t}{dz_{t-s}^\top} = \sum_{j\geq 1} \frac{\partial q_t^{(d,r)}}{\partial \mathbb{E}_t[r_{t+j}]^\top} \frac{\partial \mathbb{E}_t[r_{t+j}]}{\partial z_{t-s}^\top}.$$

Define the pass-through matrices

$$A_j \equiv -\frac{\partial \mathbb{E}_t[r_{t+j}]}{\partial p_t^\top},$$

so that $\partial \mathbb{E}_t[r_{t+j}]/\partial z_{t-s}^\top = -A_j\,(dp_t/dz_{t-s}^\top)$. Substituting into (5) gives the aggregation

$$\eta_{t,s}^{\text{DIE}} = \sum_{j\geq 1} \frac{\partial q_t^{(d,r)}}{\partial \mathbb{E}_t[r_{t+j}]^\top} A_j.$$

DIE therefore cumulates marginal return sensitivities across horizons, weighted by price-to-return pass-through. It captures portfolio responses to persistent price movements that systematically revalue long-horizon expected returns, rather than to transitory discount-rate shifts or pure cash-flow news.

3 Elasticity as a Function of Key Variables

In the preceding discussion, we noted that dynamic settings appear to give rise to multiple definitions of demand elasticity, while the quantities estimated in practice have an additional wedge terms that brings even more complexity. At first glance, estimating demand elasticities in the presence of dynamics thus seems difficult and perhaps hopeless. In this section, we show that—despite persistent outside investment—a dynamic return instrumented elasticity (DIE) can be both identified and interpreted as the appropriate measure linking flows to prices. The key qualification is that the resulting price multiplier is context-dependent: applying it to counterfactual scenarios implicitly assumes that the counterfactual environment mirrors the estimation setting. For instance, the persistence of the counterfactual investment flow must resemble that of the observed flow used to estimate the elasticity.

This section develops a tractable linear equilibrium model in levels. We follow Haddad, He, Huebner, Kondor, and Loualiche (2025a) by referring to the derivative of shares demanded with respect to prices, rather than the derivative of log shares with respect to log prices, as an elasticity despite the fact that we are using levels and not logs. Working in levels ensures that both instrumented demand elasticities and price multipliers remain linear objects. We use this framework to define and compare the Static, Immediate, and Dynamic Instrumented Elasticities and to derive closed-form expressions for the corresponding equilibrium price multipliers.

3.1 Setup

Let $\beta = (R^f)^{-1} \in (0,1)$ denote the pure discount factor, where R^f is the gross risk-free rate. There are N assets. A representative arbitrageur is mean-variance each period with risk aversion $\gamma > 0$ and a time-invariant dollar covariance matrix $\Sigma \succ 0$. Let $Q_t \in \mathbb{R}^N$ denote the arbitrageur's demand (in shares) and $Z_t \in \mathbb{R}^N$ denote exogenous outside holdings. Shares outstanding are normalized to $\mathbf{1}_N$, so market clearing is

$$Q_t + Z_t = \mathbf{1}_N. \tag{10}$$

Let $P_t \in \mathbb{R}^N$ denote prices and $D_t \in \mathbb{R}^N$ denote dividends. Define the one-period *dollar* excess return as:

$$R_{t+1} \equiv P_{t+1} + D_{t+1} - R^f P_t.$$

Under mean-variance preferences, demand is

$$Q_t = \frac{1}{\gamma} \Sigma^{-1} \mathbb{E}_t [R_{t+1}]. \tag{11}$$

As in Section 2, this demand function can be expressed equivalently using alternative state variables:

$$
\begin{aligned}
Q^{(p,d)}(P_t, \mathbb{E}_t P_{t+1}, \mathbb{E}_t D_{t+1}) &= \frac{1}{\gamma}\,\Sigma^{-1}\left(\mathbb{E}_t[P_{t+1}] + \mathbb{E}_t[D_{t+1}] - R^f\,P_t\right)\\
Q^{(p,r)}(P_t, \mathbb{E}_t R_{t+1}) = Q^{(p,r)}(\mathbb{E}_t R_{t+1}) &= \frac{1}{\gamma}\,\Sigma^{-1}\mathbb{E}_t[R_{t+1}],\\
Q^{(d,r)}(\mathbb{E}_t D_{t+1}, \mathbb{E}_t R_{t+1}) = Q^{(d,r)}(\mathbb{E}_t R_{t+1}) &= \frac{1}{\gamma}\,\Sigma^{-1}\mathbb{E}_t[R_{t+1}],
\end{aligned}
$$

The latter two representations depend only on expected returns: holding $\mathbb{E}_t[R_{t+1}]$ fixed, neither prices nor dividends affect demand. For example,

$$
\frac{\partial Q_t^{(p,r)}}{\partial P_t^\top} = 0.
$$

The elasticities implied by each representation are:

$$
\begin{aligned}
-\frac{\partial Q_t^{(p,d)}}{\partial P_t^\top} &= \frac{1}{\gamma\beta}\,\Sigma^{-1}, \qquad (12)\\
-\frac{\partial Q_t^{(p,r)}}{\partial P_t^\top} = -\frac{\partial Q_t^{(d,r)}}{\partial P_t^\top} &= 0_{N\times N}.
\end{aligned}
$$

Thus, only the price-dividend representation, $Q^{(p,d)}$, exhibits a nonzero price elasticity, reflecting the fact that, under mean-variance demand, prices matter only through their impact on expected returns.

3.2 Equilibrium

Combining (10) and (11) gives equilibrium expected excess returns:

$$
\mathbb{E}_t[R_{t+1}] = \gamma\,\Sigma\,(\mathbf{1}_N - Z_t). \qquad (13)
$$

Substituting into the pricing recursion yields

$$P_t = \beta\, \mathbb{E}_t[P_{t+1}] + \beta\, \mathbb{E}_t[D_{t+1}] - \beta\, \gamma\, \Sigma\, (\mathbf{1}_N - Z_t). \tag{14}$$

Iterating (14) forward gives, for any finite J,

$$P_t = \sum_{j=0}^{J} \beta^{j+1} \mathbb{E}_t[D_{t+j+1}] - \gamma\, \Sigma \sum_{j=0}^{J} \beta^{j+1} \big(\mathbf{1}_N - \mathbb{E}_t[Z_{t+j}]\big) + \beta^{J+1} \mathbb{E}_t[P_{t+J+1}].$$

Under the assumptions that discounted dividends are finite, i.e. $|| \sum_{j=0}^{\infty} \beta^{j+1} \mathbb{E}_t[D_{t+j+1}]|| < \infty$, and the standard transversality condition:

$$\lim_{J \to \infty} \beta^{J+1} \mathbb{E}_t[\|P_{t+J+1}\|] = 0,$$

the forward solution for prices is

$$P_t = \sum_{j=0}^{\infty} \beta^{j+1} \mathbb{E}_t[D_{t+j+1}] - \gamma\, \Sigma \sum_{j=0}^{\infty} \beta^{j+1} \big(\mathbf{1}_N - \mathbb{E}_t[Z_{t+j}]\big). \tag{15}$$

Equation (15) shows that prices are the discounted present value of expected dividends minus the discounted expected excess holdings of outside investors. The second term captures how expected future imbalances Z_{t+j} influence today's price level through risk premia.

Define $\Delta P_t \equiv P_t - P_{t-1}$ as the price change, which equals

$$\begin{aligned}\Delta P_t = &\sum_{j=0}^{\infty} \beta^{j+1} \big(\mathbb{E}_t[D_{t+j+1}] - \mathbb{E}_{t-1}[D_{t+j+1}]\big) \; + \; \gamma\, \Sigma \sum_{j=0}^{\infty} \beta^{j+1} \big(\mathbb{E}_t[Z_{t+j}] - \mathbb{E}_{t-1}[Z_{t+j}]\big) \\ &+ \; \beta\big(\mathbb{E}_{t-1}[D_{t+1}] - \mathbb{E}_{t-1}[D_t]\big) \; + \; \gamma\, \Sigma\, \beta\big(\mathbb{E}_{t-1}[Z_t] - Z_{t-1}\big).\end{aligned}$$

This decomposes price changes into four components: revisions in expected future dividends, revisions in expected future outside positions, and two boundary adjustments—one reflecting that the current dividend has been paid, and another reflecting that the current flow has already taken place.

3.3 Noisy flows

Assume outside positions follow a geometric decay with (componentwise) persistence $\rho \in [0,1)$, and we decompose innovations into an *expected* component and an *unexpected (surprise)* component:

$$Z_{t+1} = \rho\, Z_t + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} = \varepsilon^{\text{exp}}_{t,t+1} + \varepsilon^{\text{surp}}_{t+1}, \qquad \mathbb{E}_t[\varepsilon^{\text{surp}}_{t+1}] = 0.$$

If $\rho = 0$, this means the noisy investor buys (sells) one period, and immediately sells (buys) the next period. If $\rho > 0$, then the noisy investor starts to immedaitely divest the next period, but does this following a simple geometric divestment rule.

Then the conditional path is

$$\mathbb{E}_t[Z_{t+j}] = \rho^{\,j} Z_t + \sum_{k=1}^{j} \rho^{\,j-k}\, \varepsilon^{\text{exp}}_{t,t+k}. \tag{16}$$

3.4 Closed-form equilibrium prices

Plugging (16) into (15) and summing the geometric series (using $|\beta\rho| < 1$) yields

$$\begin{aligned} P_t &= \sum_{j=0}^{\infty} \beta^{j+1} \mathbb{E}_t[D_{t+j+1}] - \gamma\,\Sigma \left[\sum_{j=0}^{\infty} \beta^{j+1} - \sum_{j=0}^{\infty} \beta^{j+1} \mathbb{E}_t[Z_{t+j}] \right] \\ &= \sum_{j=0}^{\infty} \beta^{j+1} \mathbb{E}_t[D_{t+j+1}] - \gamma\,\Sigma \left[\frac{\beta}{1-\beta} \mathbf{1}_N - \frac{\beta}{1-\beta\rho} Z_t - \frac{1}{1-\beta\rho} \sum_{k=1}^{\infty} \beta^{k+1} \varepsilon^{\exp}_{t,t+k} \right]. \end{aligned} \tag{17}$$

The price change is given by:

$$\begin{aligned} \Delta P_t = \sum_{j=0}^{\infty} \beta^{j+1} \big(\mathbb{E}_t[D_{t+j+1}] - \mathbb{E}_{t-1}[D_{t+j+1}] \big) + \beta \big(\mathbb{E}_{t-1}[D_{t+1}] - \mathbb{E}_{t-1}[D_t] \big) \\ + \frac{\gamma\,\Sigma}{1-\beta\rho} \left\{ \beta\,\Delta Z_t + \beta^2 \big(\varepsilon^{\exp}_{t,t+1} - \varepsilon^{\exp}_{t-1,t} \big) + \sum_{k=2}^{\infty} \beta^{k+1} \big(\varepsilon^{\exp}_{t,t+k} - \varepsilon^{\exp}_{t-1,t-1+k} \big) \right\}, \end{aligned} \tag{18}$$

where $\Delta Z_t \equiv Z_t - Z_{t-1}$. This decomposes price changes in dividend news revisions, changes in flows, the expected flow from the prior period, and revisions of expected flows.

3.5 What kind of instrumented elasticity can we measure?

Z_t cannot be used to measure SIE. To see this, we can calculate the following from equation (13):

$$\frac{d\,\mathbb{E}_t R_{t+1}}{dZ_t^{\top}} \;=\; -\gamma\,\Sigma \;\neq\; 0.$$

Thus any instrument proportional to Z_t necessarily changes expected returns on impact, so it cannot hold expected returns fixed. Hence Z_t cannot identify a Static Instrumented Elasticity (SIE).

It is also the case that Z_t cannot be used to measure IIE. To see this, note that for $j \geq 1$,

$$\frac{d\, \mathbb{E}_t R_{t+1+j}}{dZ_t^\top} \;=\; -\gamma\, \Sigma\, \frac{d\, \mathbb{E}_t Z_{t+j}}{dZ_t^\top} \;=\; -\gamma\, \Sigma\, \rho^{\,j} \;\neq\; 0 \quad \text{if } \rho > 0.$$

In particular, for all $j \geq 2$ this derivative is nonzero (and for $j = 1$ as well), implying that moving $\mathbb{E}_t R_{t+1}$ with Z_t *also* moves $\mathbb{E}_t R_{t+1+j}$ for $j \geq 2$. Therefore Z_t cannot deliver a pure Immediate Instrumented Elasticity (IIE) unless at the knife-edge $\rho = 0$.

Thus, in this equilibrium model, there are no shocks that allows us to measure either the SIE or IIE. If an economist only has access to nonfundmental and persistent investment flows, as is the situation here, then SIE or IIE may not be possible to identify.

Importantly, DIE is identifiable with Z_t. From equation (17) we can calculate the following derivative:

$$\frac{dP_t}{dZ_t^\top} \;=\; \gamma\, \Sigma\, \frac{\beta}{1 - \beta\rho}.$$

Market clearing gives $Q_t = \mathbf{1}_N - Z_t$, with contemporaneous flows ($s = 0$ lags), we can write:

$$\frac{\partial Q_t}{\partial Z_t^\top} = -I \qquad \Longrightarrow \qquad \eta^{\text{DIE}} \;\equiv\; =\; -\frac{\partial Q_t}{\partial Z_t^\top} \left(\frac{\partial P_t}{\partial Z_t^\top} \right)^{-1} \;=\; \frac{1 - \beta\rho}{\gamma\, \beta}\, \Sigma^{-1}. \tag{19}$$

Thus, while flow instruments cannot identify SIE or IIE in general, they *do* identify the Dynamic Instrumented Elasticity (DIE).

3.6 Price multipliers

Let $\Delta P_t \equiv P_t - P_{t-1}$ denote the realized price change between $t - 1$ and t. We define the *contemporaneous price multiplier* as the sensitivity of this realized change to the period-t flow shock:

$$M_t \;\equiv\; \frac{d\,\Delta P_t}{d\,Z_t^\top}. \tag{20}$$

Using the decomposition in (18), the coefficient on Z_t is

$$M_t \;=\; \frac{\gamma\,\beta}{1-\beta\rho}\,\Sigma, \tag{21}$$

Thus, if the outside investor buys 1% of shares outstanding, then prices are higher by $0.01(\gamma\beta)/(1-\beta\rho)\Sigma$ dollars.

3.7 Link between $Q^{(p,d)}$ elasticity, the DIE, and the price multiplier

For ease of comparison, from equations (12), (19), and (21) we can write:

$$-\frac{\partial Q^{(p,d)}}{\partial P_t^\top} = \frac{1}{\gamma\beta}\,\Sigma^{-1}, \quad \eta^{\text{DIE}} = (1-\beta\rho)\left(-\frac{\partial Q^{(p,d)}}{\partial P_t^\top}\right) = \frac{1-\beta\rho}{\gamma\beta}\,\Sigma^{-1}, \quad M_t = (\eta^{\text{DIE}})^{-1} = \frac{\gamma\beta}{1-\beta\rho}\,\Sigma.$$

Thus, in words, the multiplier matrix is the inverse of the DIE matrix, and not the inverse of the elasticity that holds fixed future expectations of cash flows and prices, $-\partial Q_t^{(p,d)}/\partial P_t^\top$.

The equilibrium relationships above yield two clear insights. When flows are persistent, changes in prices necessarily alter the entire expected-return path. In that setting, instruments based on flows cannot identify either the static or the immediate instrumented elasticity, because any shift in flows moves all future returns. If one nevertheless uses the dynamic elasticity to approximate the elasticity that holds fixed future prices and dividends, the result is systematically attenuated by the persistence factor $(1-\beta\rho)$. The higher the persistence, the further the estimated elasticity drifts from the true ceteris paribus object.

The second insight is constructive. For most applications, the object of interest is not the elasticity that holds expectations fixed, but the mapping from flows to prices—the price multiplier. In this model, the dynamic instrumented elasticity is exactly the inverse of that multiplier. Estimating the DIE and inverting it therefore delivers the correct counterfactual price impact of a flow with persistence (ρ). The elasticity that holds future prices and dividends fixed, by contrast, does not correspond to the equilibrium adjustment that actually governs prices.

The broader implication is straightforward. In economies with persistent investment flows, the DIE is not merely a convenient empirical construct—it is the theoretically correct elasticity for linking observed flows to prices. Attempts to estimate a “static" elasticity from such data will be both infeasible and economically misleading, while the DIE provides the right quantitative object for understanding how capital reallocations move market prices.

3.8 Determinants of Price Multipliers

Equation (21) gives rise to several comparative-static results:

Proposition 1 (Return Variance)**.** *The price multiplier is larger if the variance of the stock is larger (i.e., the corresponding diagonal element of Σ is larger).*

Proposition 2 (Flow Persistence)**.** *The price multiplier is larger if the flow that induces the price change is more persistent (i.e., if ρ is larger).*

Systematic versus Idiosyncratic Directions. Using equation (18) and the definition of the contemporaneous price multiplier in (20), we can write the linear law of motion for prices as

$$\Delta P_t \ = \ \frac{\gamma\beta}{1-\beta\rho}\,\Sigma\,\Delta Z_t \ + \ \zeta_t,$$

where ζ_t collects all remaining terms in (18).

Consider a directional flow $Z_t = \omega v$, where ω is a scalar and v is a unit eigenvector of Σ with eigenvalue λ. Differentiating with respect to ω yields the directional price multiplier:

$$M_t(v) \;\equiv\; \frac{\partial \Delta P_t}{\partial \omega} \;=\; \frac{\gamma\,\beta}{1-\beta\rho}\,\lambda\,v.$$

Because v is normalized, the magnitude of the price response is determined entirely by the eigenvalue λ. Flows aligned with eigenvectors associated with larger λ—that is, more systematic directions of risk—generate proportionally larger price movements.

Proposition 3 (Systematic versus Idiosyncratic)**.** *The magnitude of the directional multiplier, $\|M_t(v)\|$, increases monotonically with the corresponding eigenvalue λ of Σ. Hence, systematic flows (aligned with high-λ eigenvectors) move prices more than idiosyncratic flows of equal size.*

Information Timing. Finally, we consider the timing of information. Define the following:

$$M_t^{\text{surp}} \;\equiv\; \frac{\partial\,\Delta P_t}{\partial(\varepsilon_t^{\text{surp}})^\top}, \qquad M_t^{\text{exp}} \;\equiv\; \frac{\partial\,\Delta P_t}{\partial(\varepsilon_{t-1,t}^{\text{exp}})^\top},$$

which is the price multiplier for surprise flows and expected flows respectively. This leads to our next proposition:

Proposition 4 (Expected versus Surprise Flow)**.** *Holding fixed beliefs at t about flows strictly after t (i.e., $\varepsilon_{t,t+k}^{\text{exp}} = \varepsilon_{t-1,t+k}^{\text{exp}}$ for all $k \geq 1$),*

$$M_t^{\text{surp}} \;=\; \frac{\gamma\beta}{1-\beta\rho}\,\Sigma, \qquad M_t^{\text{exp}} \;=\; \frac{\gamma\beta(1-\beta)}{1-\beta\rho}\,\Sigma \;=\; (1-\beta)\,M_t^{\text{surp}}.$$

To calculate these, just plug $\Delta Z_t = (\rho-1)Z_{t-1} + \varepsilon_{t-1,t}^{\text{exp}} + \varepsilon_t^{\text{surp}}$ into (18), and calculate the derivatives.

In words, when multipliers are defined with respect to ΔP_t, the expected-flow multiplier is a fraction $(1-\beta)$ of the surprise multiplier. If interest rates are low, then β is close to one and $(1-\beta)$ is close to zero, so expected (anticipated) flows move realized prices much less than equally sized surprise flows.

While these results are written in terms of price multipliers, the parallel statements in terms of the DIE follow immediately because $M_t = (\eta^{\text{DIE}})^{-1}$. The arbitrageur is less elastic (i.e., the multiplier is larger and DIE is smaller) for stocks with larger variance and for price shocks induced by more persistent flows. Arbitrageurs are less elastic to more systematic shocks than to idiosyncratic shocks. Finally, because $M_t^{\text{exp}} = (1-\beta)M_t^{\text{surp}}$, arbitrageurs are *more* elastic to anticipated flows (smaller multiplier, larger DIE) and *less* elastic to surprise flows.

In summary, this section links investor demand elasticities to price multipliers in a tractable linear equilibrium. The price multiplier for realized price changes is $M_t = \frac{\gamma\beta}{1-\beta\rho}\Sigma$, which scales with risk (via Σ), persistence (via ρ), and systematic exposure (via eigenvalues of Σ). Decomposing flows into expected and surprise components shows that the expected-flow multiplier is dampened by $(1-\beta)$ relative to the surprise multiplier. Because $M_t = (\eta^{\text{DIE}})^{-1}$, these comparative statics map one-for-one into properties of the Dynamic Instrumented Elasticity.

4 Empirical Results

In this section, we use two demand shocks in the previous literature to test the predictions in Section 3. We find that price multipliers indeed vary as predicted, which suggests that there are empirical regularities in price impacts that are generalizable.

4.1 Data

We use two measures of demand which prior papers have argued to reflect trading that is unrelated to cash flow information. For more details about their construction and arguments for why they are unrelated to cash flows, please see the original papers. In our main specification, we compute weekly versions of these variables.

Flow-Induced Trading (FIT) Our first demand measure is flow-induced trading (FIT) introduced in Lou (2012), which captures the non-discretionary stock-level trading of mutual funds and ETFs in response to flows.[3] The measure relies on a specfiic mechanism: in response to inflows (outflows), funds tend to scale up (down) their pre-existing holdings in a non-discretionary fashion (Coval and Stafford, 2007; Lou, 2012). For example, if Apple's existing weight is 5% in a fund's portfolio, a $1 inflow (outflow) induces the fund allocate about five more (fewer) cents to Apple. This behavior is true not only for index funds, but also, on average, for active mutual funds and ETFs (see Figure A4 in Li, 2022). Frazzini and Lamont (2008) and Lou (2012) show that the price impact of FIT reverts over time, which suggests that such trades are not motivated by superior cash flow information.

While other researchers have used monthly or quarterly FIT (Lou, 2012; Ben-David, Li, Rossi, and Song, 2022, e.g.) To obtain higher frequency measures, we use the weekly fund flows from TrimTabs and merge it with 13F fund holdings data from Thomson Reuters. We compute FIT for stock i in week t as:

$$\mathrm{FIT}_{i,t} = \frac{\sum_{\text{fund } j} \mathrm{SharesHeld}_{i,j,t-1} \cdot \mathrm{flow}_{j,t}}{\mathrm{SharesOutstanding}_{n,t-1}},$$

where $\mathrm{flow}_{j,t}$ is the net flow in fund j as a fraction of its lagged assets, $\mathrm{ShareHeld}_{i,j,t-1}$ is the lagged

[3]While not initially intended as a demand instrument, it has been used as such by subsequent work (e.g. Li, 2022; Chaudhry, 2025; Van der Beck, 2025).

number of shares of stock i held by fund j, and the denominator is the lagged number of shares outstanding for stock i. Normalizing by shares outstanding ensures that when we regress stock returns on FIT, the resulting coefficient is in "price multiplier" units (Gabaix and Koijen, 2022).

Order flow imbalance (OFI). Our second demand shock is the Lee-Ready signed order flow imbalance (OFI). This measure takes all trades in the U.S. stock market and classifies them as buyer- or seller-initiated using the Lee and Ready (1991) algorithm. We download daily Lee-Ready signed OFI data from WRDS intraday indicators, aggregate to weekly frequencies, and then normalize by lagged shares outstanding.[4] The OFI measure has been used heavily in the microstructure literature to study price impact at daily or higher frequencies. Recently, Li and Lin (2023) argue that OFI can also be useful for studying price impact at "asset pricing frequencies" (i.e. monthly or lower).

The main benefit of OFI is that it provides much more variation than FIT. While FIT only captures a subset of trades by mutual funds and ETFs, OFI captures all aggressively executed trading in the U.S. stock market and thus have much more variation. The main drawback is the relative lack of knowledge about the drivers of OFI. Li and Lin (2023) conduct extensive tests and do not find evidence that OFI is related to various measures of cash-flow news. However, they also admit that, due to the difficulty of measuring news, one cannot be fully certain that OFI is unrelated to cash flow news.

Table 2 summarizes our variables. FIT, which is available starting from 2007 due to availability of TrimTabs flow data, has a mean around zero and a standard deviation of 0.08%. This means that the typical FIT realization for a stock in a week amounts to buying or selling 0.08% of shares outstanding of a stock. OFI, which is available from 1993, has more variation: the weekly standard deviation is 0.68% which is almost an order of magnitude larger than that of FIT.

[4] OFI only captures all trades that are executed *aggressively*. Many sophisticated institutional investors tend to execute trades slowly and passively to reduce price impact; OFI does not capture those trades.

				Percentiles						
	Obs	Mean	StDev	1%	5%	25%	50%	75%	95%	99%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Return	4,486	0.25	8.54	-22.24	-11.71	-3.30	-0.00	3.23	12.82	28.57
FIT	3,555	0.01	0.08	-0.23	-0.13	-0.03	0.00	0.04	0.15	0.27
OFI	4,431	-0.05	0.68	-2.63	-0.93	-0.20	-0.02	0.15	0.80	2.03
Idiosyncratic vol (63 day)	4,492	3.22	2.20	0.71	0.98	1.69	2.61	4.08	7.57	11.23
Idiosyncratic vol (252 day)	4,492	3.28	2.03	0.79	1.09	1.81	2.74	4.20	7.33	10.15

Table 2. Summary Statistics.
We report summary statistics of the main weekly variables in percent. The demand measures are expressed as fraction of shares outstanding. FIT is flow-induced trading as defined in Lou (2012) and computed using weekly flow data from TrimTabs, and OFI is the order flow imbalance measure used in Li and Lin (2023). The sample for return and OFI spans 1993 to 2022, while the sample for FIT starts from 2007 due to availability of weekly fund flow data. Column (1) reports the average number of stocks per period. Idiosyncratic return is defined as stock returns purged of market and industry components, and the reported volatility is for daily returns.

4.2 Multiplier Increases with Return Variance

We test the prediction that price impacts increases with return variance (Proposition 1). For each stock i and week t, we use the previous 21 (or 63) trading days — which corresonds to roughly a month (or a quarter) — to compute the rolling idiosyncratic return volatility, where idiosyncratic return are obtained via time-series regressions of stock returns on market and Fama-French industry returns. We then estimate panel regressions of stock returns on flows interacted with lagged volatilties:

$$r_{i,t} = b_0 \cdot f_{i,t} + b_1 \cdot f_{i,t} \times \sigma_{i,t-1} + \tau_t + \tau_i \epsilon_{i,t}$$

where the controls include time and stock fixed effects. We cluster standard errors by time and stock. We aim to test whether $b_1 > 0$, that is, whether multipliers are larger when return volatilities are higher.

The regression results are reported in Table 3. Columns (1)-(2) show results for FIT where they differ by the lookback window for computing volatilities. The results indicate that multipliers indeed increase with return volatility. For each 1 standard deviation increase in volatility (0.34), price multiplier increases by around $0.34 \times 6.92 \approx 2.35$ to $0.34 \times 8.75 \approx 2.98$, which is sizeable when compared to the multiplier for an average stock (e.g., per column (1), $1.32 + 0.48 \times 6.92 \approx 4.64$). The results are also statistically significant. Corresponding OFI results in columns (5)-(6) are qualitatively consistent.

To examine whether the effect is monotonic in volatilities, we also sort stocks by $\sigma_{i,t-1}$ in each cross-section into quintiles and interact $f_{i,t}$ with those quintile indiators. The results are reported in columns (3)-(4) and (7)-(8). The conclusion is similar and demonstrates that the effect is monotonic. Overall, the results are consistent with the prediction in Proposition 1.

Dependent variable: stock return $r_{i,t}$								
$f_{i,t}$ =	FIT				OFI			
Vol window	1m	1q	1m	1q	1m	1q	1m	1q
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$f_{i,t}$	1.32**	0.60	1.83***	1.84***	1.07***	0.68***	1.22***	1.18***
	(0.66)	(0.84)	(0.44)	(0.47)	(0.04)	(0.05)	(0.03)	(0.03)
$f_{i,t} \times \sigma_{i,t-1}$	6.92***	8.75***			2.43***	3.02***		
	(1.68)	(2.17)			(0.09)	(0.10)		
$f_{i,t} \times I_{\sigma_{i,t-1}\ \text{bin}=2}$			1.04***	0.79***			0.50***	0.41***
			(0.24)	(0.25)			(0.02)	(0.02)
$f_{i,t} \times I_{\sigma_{i,t-1}\ \text{bin}=3}$			1.86***	1.87***			1.03***	0.87***
			(0.37)	(0.42)			(0.03)	(0.03)
$f_{i,t} \times I_{\sigma_{i,t-1}\ \text{bin}=4}$			3.26***	3.20***			1.65***	1.56***
			(0.56)	(0.57)			(0.04)	(0.04)
$f_{i,t} \times I_{\sigma_{i,t-1}\ \text{bin}=5}$			5.19***	5.91***			2.52***	2.80***
			(0.93)	(1.11)			(0.06)	(0.06)
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Stock FE	Y	Y	Y	Y	Y	Y	Y	Y
Obs	2,249,682	2,236,975	2,249,682	2,236,975	7,424,584	7,333,328	7,424,584	7,333,328
R^2	0.177	0.178	0.177	0.178	0.156	0.158	0.153	0.155
Within R^2	0.002	0.002	0.002	0.002	0.058	0.060	0.055	0.057

Table 3. Price Impact by Return Volatility.
We estimate panel regressions of weekly stock return $r_{i,t}$ on contemporaneous flow $f_{i,t}$ and its interactions with lagged annualized idiosyncratic volatility. Column (1)-(4) and (5)-(8) report results for FIT and OFI, respectively. Columns (1)-(2) and (5)-(6) estimate effects of flows interacted with the value of lagged volatiltiy, while the other columns estimate effects of flows interacted with quintile indicators of lagged return volatility sorted in each cross-section. Odd (even) columns estimate return volatility using 1 month (1 quarter) lookback windows. All regressions control for time and stock fixed effects. Standard errors are clustered by time and stock, and they are reported in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.3 Multiplier Increases with Flow Persistence

Proposition 2 in Section 3 states that the price multiplier scales up with the persistence of flows. To test this, we measure the persistence of FIT and OFI positions, once the expected component ($\varepsilon^{\text{exp}}_{t,t+k}$) is extracted. We then estimate a regression to see if price multipliers increases with persistence.

Constructing $Z_{i,t}^{\textbf{surp}}$ and ρ_i. Let $f_{i,t}$ be the OFI or FIT flow at time t. We estimate the position level by summing up all flows from the beginning of the data sample:

$$Z_{i,t} = \sum_{\tau \leq t} f_{i,\tau}$$

To remove the expected component, we first detrend the cumulative position series ($Z_{i,t}$) by stock to isolate the surprise component. Specifically, starting from $Z_{i,t}$ we obtain $Z_{i,t}^{\text{det}}$ by either a cubic LSQ spline (*Spline*) or a flexible Hodrick–Prescott filter (*HP*). We denote the detrended component of $Z_{i,t}$ as $Z_{i,t}^{\text{det}}$. We then model the detrended positions for each stock i as a stationary AR(1) in levels with a residual $Z_{i,t}^{\text{surp}}$:

$$Z_{i,t}^{\text{det}} = \rho_i \, Z_{i,t-1}^{\text{det}} + Z_{i,t}^{\text{surp}}.$$

This is estimated by OLS without an intercept on all valid $(t-1,t)$ pairs for stock i. Given $\widehat{\rho}_i$, the *position surprise* $Z_{i,t}^{\text{surp}}$ isolates the unanticipated component of the detrended position at t. Intuitively, $\widehat{\rho}_i \, Z_{i,t-1}^{\text{det}}$ is the expected continuation of the position based on its own persistence, while $Z_{i,t}^{\text{surp}}$ is the innovation that theory predicts should move prices on impact. These objects—$Z_{i,t}^{\text{surp}}$ and $\widehat{\rho}_i$—enter the regression in interaction to test whether price impact scales with persistence.

Testing whether price multipliers are a increasing function of persistence. To bring this prediction to the data in a transparent and model-light way, we use a reduced-form parameterization in which the (conditional) multiplier is allowed to depend linearly on ρ_i, written as $(\beta_0 + \beta_1 \rho_i)$, where β_0 and β_1 are freely estimated parameters. We estimate the regression:

$$r_{i,t} \;=\; \alpha_t \;+\; \beta_0 \, Z_{i,t}^{\text{surp}} \;+\; \beta_1 \left(Z_{i,t}^{\text{surp}} \times \rho_i \right) \;+\; \varepsilon_{i,t},$$

	FIT (Spline)	FIT (HP)	OFI (Spline)	OFI (HP)
$Z^{\text{surp}} \times \rho_i$	5.782*	27.518***	1.575***	2.324***
	(3.383)	(9.920)	(0.122)	(0.296)
Z^{surp}	4.221**	16.166***	1.979***	3.774***
	(1.722)	(3.982)	(0.072)	(0.096)
# Observations	2,245,961	2,245,961	7,455,959	7,455,959
R^2	0.004	0.001	0.039	0.012

Table 4. Price Impact and Flow Persistence.
We estimate the weekly panel regression $r_{i,t} = \alpha_t + \beta_0 Z^{\text{surp}}_{i,t} + \beta_1 (Z^{\text{surp}}_{i,t} \times \rho_i) + \varepsilon_{i,t}$, where $r_{i,t}$ is the return, $Z^{\text{surp}}_{i,t}$ is the position surprise, and ρ_i is the estimated persistence of detrended positions for stock i. All specifications absorb time fixed effects and cluster standard errors by time. The interaction term tests whether the price multiplier is increasing in flow persistence, as predicted by Proposition 2. Both FIT (flow-induced trading) and OFI (order flow imbalance) measures are analyzed using cubic spline and Hodrick-Prescott (HP) detrending. Dependent variable is the weekly stock return. All specifications absorb time FEs and cluster by time. Standard errors are reported in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

where α_t are time fixed effects and standard errors are clustered by time. The interaction term $Z^{\text{surp}}_{i,t} \times \rho_i$ tests whether the flow-induced price impact scales with persistence. We estimate this regression for both weekly FIT and OFI data, using both spline and HP detrending to ensure that the results are not sensitive to the method used to remove low-frequency trends from positions.

Results. Table 4 reports the results of estimating equation (4.3) for both weekly FIT and weekly OFI flows using two detrending methods (Spline and HP). Across all specifications, the interaction term $Z^{\text{surp}}_{i,t} \times \rho_i$ is positive and statistically significant, consistent with the model's prediction that more persistent flows generate larger price impacts.

For FIT flows, the coefficients on $Z^{\text{surp}}_{i,t} \times \rho_i$ are large in magnitude—5.8 under spline detrending and 27.5 under the HP filter—and significant at the 10% and 1% levels, respectively. OFI flows exhibit the same qualitative pattern with smaller magnitudes, but the interaction coefficients are strongly statistically significant. Overall, the estimates indicate that price multipliers increase with flow persistence: a one–standard-deviation increase in ρ_i roughly doubles the effect of a flow surprise on returns.

4.4 Systematic Flows have Larger Multipliers Than Idiosyncratic Flows

To test the prediction in Proposition 3 that systematic demand variation has larger price impact, we decompose flows into sysetmatic and idiosyncratic components. Concretely, we use cross-sectional regressions to project stock-level flows ($f_{i,t}$) onto 14 commonly used stock characteristics downloaded from Chen and Zimmermann (2021) and Fama-French 12 industry dummy indicators.[5] For each time period t, we estimate realized demand factors $\{g_{k,t}\}_{k=1}^{K}$ via cross-sectional regressions

$$f_{i,t} = \underbrace{a_t + \sum_{k=1}^{K} b_{i,k,t-1} \cdot \hat{g}_{k,t}}_{\equiv f_{i,t}^{\text{systematic}}} + \underbrace{\hat{u}_{i,t}}_{\equiv f_{i,t}^{\text{idiosyncratic}}},$$

which we then use to decomopse flows into systematic and idiosyncratic comopnents. We then estimate price multipliers using panel regressions with time fixed effects:

$$r_{i,t} = M_{\text{systematic}} \cdot f_{i,t}^{\text{systematic}} + M_{\text{idiosyncratic}} \cdot f_{i,t}^{\text{idiosyncratic}} + \tau_t + v_{i,t}$$

and cluster standard errors by time and stock. Proposition 3 implies that $M_{\text{systematic}} > M_{\text{idiosyncratic}}$.

The regressions results are reported in Table 5. Columns (1) and (2) report results based on the full sample for FIT and OFI, respectively. For both of them, we find that the systematic component-multiplier is significantly higher than the idiosyncratic component-multiplier, as indicated by the last row in the table. In columns (3) to (5), we further split OFI — which has more statistical power than FIT — into three subperiods and estimate results separately. We find that the conclusion is

[5] The characteristics include accruals, asset growth, beta, book-to-market ratio, gross profitability, industry momentum, intermediate momentum, 1-year issuance, 5-year issuance, momentum, seasonal momentum, net operating assets, realized volatility, and size. Following common practice (e.g. Kelly, Pruitt, and Su, 2019), we transform each characteristics into uniform distributions from -0.5 to 0.5 by rank in each cross-section.

Dependent variable: stock return $r_{i,t}$					
$f_{i,t}$ =	FIT	OFI			
Sample	Full	Full	1993-2002	2003-2012	2013 - 2022
	(1)	(2)	(3)	(4)	(5)
$f_{i,t}^{\text{systematic}}$	6.39***	4.40***	5.29***	3.36***	3.94***
	(1.18)	(0.25)	(0.45)	(0.25)	(0.38)
$f_{i,t}^{\text{idiosyncratic}}$	3.49***	2.55***	2.92***	2.22***	1.94***
	(0.35)	(0.04)	(0.06)	(0.06)	(0.05)
Time fixed effect	Y	Y	Y	Y	Y
Obs	2,253,935	7,475,646	3,156,053	2,402,155	1,917,438
R^2	0.176	0.147	0.122	0.183	0.170
R^2 of $f_{i,t}$	0.002	0.050	0.063	0.039	0.031
$f_{i,t}^{\text{systematic}} - f_{i,t}^{\text{idiosyncratic}}$	2.90***	1.85***	2.37***	1.14***	2.00***
	(1.07)	(0.24)	(0.44)	(0.25)	(0.38)

Table 5. Price Impact for Systematic versus Idiosyncratic Flows.
For each demand measure, we decompose demand into an aggregate component ($f_{i,t}^{\text{systematic}}$) that corresponds to industry- and stock characteristics-related demand, and a residual ($f_{i,t}^{\text{idiosyncratic}}$). We then use panel regressions to estimate the price multiplier associated with these demand components, controlling for time fixed effect and clustering standard errors by time and stock. Column (1) reports results for FIT. Columns (2)-(5) report results for OFI, with column (2) using the full sample and the other columns using subsamples. The last row of the table reports the multiplier difference with standard errors computed using the Delta method. Standard errors are reported in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

robust across subperiods. Overall, the evidence is consistent with the prediction that systematic demand leads to larger price impact.[6]

[6] Li and Lin (2023) arrive at a similar conclusion using different specifications.

4.5 More Surprising Flows have Larger Multipliers

We now test Proposition 4 which states that surprise flows have larger price impact than expected flows. In theory, a flow is either expected or unexpected. In practice, if investors have heterogeneous information sets, then there can also be intensive margin variation: some flows may be expected to a higher or lower fraction of investors. Therefore, we decompose realized flow $f_{i,t}$ into multiple components into three components with decreasing degree of "surprisingness":

$$f_{i,t} = f_{i,t}^{\text{surp}} + f_{i,t}^{\text{less surp}} + f_{i,t}^{\text{predicted}}.$$

To construct these components, we fit two statistical models that predict flows using past data, and the two models differ in the amount of input variables. The more granular model, whose prediction we denote $\hat{f}_{i,t}^{\text{granular}}$, uses up to 13 weekly lags of past demand and returns (the period of approximately a quarter). The coarse model ($\hat{f}_{i,t}^{\text{coarse}}$) is based on less data: 4 weekly lags of past demand and returns (approximately a month). Then, we define the components as:

$$\begin{aligned}
f_{i,t}^{\text{surp}} &= f_{i,t} - \hat{f}_{i,t}^{\text{granular}}, \\
f_{i,t}^{\text{less surp}} &= \hat{f}_{i,t}^{\text{granular}} - \hat{f}_{i,t}^{\text{coarse}}, \\
f_{i,t}^{\text{predicted}} &= \hat{f}_{i,t}^{\text{coarse}}.
\end{aligned} \tag{22}$$

To achieve better predictive power than linear models, we use the XGBoost (Extreme Gradient Boosting) model which takes into account nonlinear interactions between input variables (Chen and Guestrin, 2016). In short, the model builds an ensemble of predicton trees, with each incremental decision tree trained on the residual prediction errors of the existing trees. This is an off-the-shelf supervised learning model commonly used in the machine learning community and we refer the reader to standard online documents for further details.[7]

To approximate the belief of investors who learn in real time, in each year y, we use data in years $y-4,\ldots,y-1$ to build a model and compute predictions on year y.[8] In unreported tests, we verify that the out-of-sample forecasts are well-calibrated, and that the model performance increases as more lagged variables are added as input variables.

Having decomposed flows as in (22), to test whether more surprising demand has larger price impact, we estimate multipliers via regression

$$r_{i,t} = b_{\text{surp}} \cdot f_{i,t}^{\text{surp}} + b_{\text{less surp}} \cdot f_{i,t}^{\text{less surp}} + b_{\text{predicted}} \cdot f_{i,t}^{\text{predicted}} + \tau_t + \text{Controls}_{i,t} + \epsilon_{i,t},$$

where τ_t denote period-specific effects and we control for 13 weekly lags of past returns to control for short-term reversal and momentum patterns. We can estimate this either using panel regression and cluster standard errors by time and stock, or using Fama-MacBeth regressions.

The results are reported in Table 6. Columns (1)-(2) report results based on FIT. As predicted by theory, $f_{i,t}^{\text{surp}}$ is associated with larger price multipliers. Panel B reports the pairwise multiplier differences. FIT-based inference has less data and smaller statistical power. While we find statistically significant differences between the coefficients of $f_{i,t}^{\text{surp}}$ and others, we do not discern a significant difference between $f_{i,t}^{\text{less surp}}$ $f_{i,t}^{\text{predicted}}$ components.

Columns (3)-(4) report results based on OFI. As discussed earlier, OFI has more variation and higher explanatory power over returns, as shown in the last row of Panel A: OFI demand can explain around 5% of weekly return variation as opposed to less than 1% in the case of FIT. Based on the panel regression in column (3), surprise demand $f_{i,t}^{\text{surp}}$ is associated with a price mutliplier

[7]For instance, see https://xgboost.readthedocs.io/en/stable/.

[8]For learning parameters, we use a max tree depth of 3, a learning rate of 0.1, and 100 rounds of boosting.

of 3.07, while the less surprising component $f_{i,t}^{\text{less surp}}$ has a multiplier of 0.45, and the predicted component $f_{i,t}^{\text{predicted}}$ has a multiplier that is indistinguishable from zero. The pairwise multiplier comparisons in Panel B are all statistically significant at the 1% level. The Fama-MacBeth results in column (4) are qualitatively similar. Overall, the results are consistent with the idea that more surprising demand have larger per-unit price impact.

5 Recovering Demand in a Dynamic Model

This section shows how reduced-form price responses to flow instruments identify the underlying intertemporal demand system. In the simple benchmark, with mean-variance demand and a persistent outside-investment process, the dynamic instrumented elasticity equals $\eta^{\text{DIE}} = \frac{1-\beta\rho}{\gamma\beta}\,\Sigma^{-1}$, whereas the contemporaneous structural object is $-\partial Q^{(p,d)}/\partial P_t^{\top} = \frac{1}{\gamma\beta}\,\Sigma^{-1}$. Observing the persistence parameter ρ therefore allows one to back out the structural marginal demand from η^{DIE}. The argument below establishes a general version of this result: once the time–series dynamics of the instrument are measured, the entire sequence of structural demand matrices can be recovered from reduced-form price impulse responses.

5.1 General linear intertemporal demand

Let $Q_t \in \mathbb{R}^N$ denote arbitrageurs' demand and $Z_t \in \mathbb{R}^N$ the outside-investor (instrument) demand. Intertemporal demand is

$$Q_t = \overline{Q} - \sum_{s=0}^{S} A_s \, \mathbb{E}_t[P_{t+s}] + F_t, \tag{23}$$

Panel A: price impact regressions				
	Dependent variable: stock return $r_{i,t}$			
$f_{i,t}$ =	FIT		OFI	
	Panel	Fama-MacBeth	Panel	Fama-MacBeth
	(1)	(2)	(3)	(4)
$f_{i,t}^{\text{surp}}$	4.09***	3.95***	3.07***	2.76***
	(0.50)	(0.24)	(0.06)	(0.03)
$f_{i,t}^{\text{less surp}}$	1.98**	1.45**	0.45***	0.68***
	(0.87)	(0.66)	(0.13)	(0.11)
$f_{i,t}^{\text{predicted}}$	2.12**	1.69*	−0.07	0.11*
	(1.07)	(0.86)	(0.10)	(0.06)
Controls	Y	Y	Y	Y
Obs	1,415,002	1,415,002	5,573,138	5,573,138
R^2	0.137	0.143	0.181	0.184
Marginal R^2 of $f_{i,t}$	0.002	0.006	0.048	0.050

Panel B: Coefficient differences				
	(1)	(2)	(3)	(4)
$f_{i,t}^{\text{surp}} - f_{i,t}^{\text{less surp}}$	2.11**	2.50***	2.61***	2.09***
	(0.92)	(0.65)	(0.15)	(0.12)
$f_{i,t}^{\text{less surp}} - f_{i,t}^{\text{predicted}}$	−0.13	−0.25	0.52***	0.56***
	(1.06)	(0.86)	(0.11)	(0.09)
$f_{i,t}^{\text{surp}} - f_{i,t}^{\text{predicted}}$	1.98*	2.26***	3.13***	2.65***
	(1.12)	(0.86)	(0.12)	(0.08)

Table 6. Price Impact for Surprise versus Expected Flows.
We use XGBoost machine learning algorithm to decompose flow $f_{i,t}$ into components that is surprising, less surprising, and predicted. We then estimate their price impacts by regressing stock returns on those components, controlling for 13 lags of past weekly returns. Columns (1)-(2) present results for FIT and columns (3)-(4) present results for OFI. Columns (1) and (3) estimate panel regressions with time fixed effects, with standard errors clustered by time and stock. Columns (2) and (4) estimate Fama-MacBeth regressions. Panel A reports the regression coefficients and its last row reports the marginal R^2 attributed to demand variables. Panel B reports pairwise coefficient comparisons where, in the case of panel regressions, the standard errors are computed using the Delta method. Standard errors are reported in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

where $A_s \in \mathbb{R}^{N\times N}$ with A_0 invertible, $P_t \in \mathbb{R}^N$ are log prices, and F_t collects fundamental shifters.

The outside-investment process admits a Wold representation

$$Z_t = \overline{Z} + \sum_{k=0}^{\infty} \Psi_{Z,k}\, \varepsilon_{Z,t-k},$$

with $\varepsilon_{Z,t}$ white noise and $\{\Psi_{Z,k}\}$ absolutely summable. Market clearing is

$$Q_t \;+\; Z_t \;=\; \mathbf{1}_N. \tag{24}$$

The matrices $\{A_s\}$ encode the structural mapping from expected future prices to current demand: A_0 captures contemporaneous elasticity and A_s for $s > 0$ capture intertemporal responses.

5.2 Equilibrium price characterization

Combining (23) and (24) yields

$$\sum_{s=0}^{S} A_s\, \mathbb{E}_t[P_{t+s}] \;=\; Y_t, \qquad Y_t \;\equiv\; Z_t \;+\; F_t \;+\; \overline{Q} \;-\; \mathbf{1}_N.$$

Let Y_t have Wold form $Y_t = \overline{Y} + \sum_{k\geq 0} \Psi_{Y,k}\, \varepsilon_{Y,t-k}$. Then equilibrium prices admit

$$P_t \;=\; \overline{P} \;+\; \sum_{k=0}^{\infty} \Psi_{p,k}\, \varepsilon_{Y,t-k}, \tag{25}$$

with $\{\Psi_{p,k}\}$ solving the linear system

$$\sum_{s=0}^{S} A_s\, \Psi_{p,k+s} \;=\; \Psi_{Y,k}, \qquad k \geq 0. \tag{26}$$

When fundamental shocks are either absent or partialled out, $\Psi_{Y,k} = \Psi_{Z,k}$, and (26) links the price impulse responses to the demand system $\{A_s\}$ and the flow dynamics $\{\Psi_{Z,k}\}$.

5.3 Recovering structural demand from reduced-form objects

Let

$$\xi_s \;\equiv\; \frac{\partial \mathbb{E}_t[P_{t+s}]}{\partial \varepsilon_{Z,t}^{\top}}, \qquad s \geq 0,$$

denote the reduced-form price impulse responses to the flow shock. Under (25) with $Y_t \equiv Z_t$, we have $\xi_s = \Psi_{p,s}$. Substituting $\Psi_{p,k} = \xi_k$ and $\Psi_{Y,k} = \Psi_{Z,k}$ into (26) gives, for all $k \geq 0$,

$$\sum_{s=0}^{S} A_s\, \xi_{k+s} \;=\; \Psi_{Z,k}. \tag{27}$$

Equation (27) is a linear convolution system in the unknowns $\{A_s\}_{s=0}^{S}$, with known right-hand side $\{\Psi_{Z,k}\}$ and known regressors $\{\xi_{k+s}\}$. Provided the flow process exhibits enough persistence so that $\{\xi_{k+s}\}$ are of full rank over a sufficient set of k's, the collection $\{A_s\}$ is point identified.

Proposition 5 (Identification and recovery of $\{A_s\}$)**.** *Suppose $\{\xi_s\}_{s\geq 0}$ and $\{\Psi_{Z,s}\}_{s\geq 0}$ are observed, with ξ_s nonsingular for at least $S+1$ distinct values of s. Then the structural demand matrices $A_0, \ldots, A_S$ are uniquely determined by the system* (27)*.*

Proof of Proposition 5. Define the formal matrix power series

$$A(z) = \sum_{s=0}^{S} A_s z^s, \qquad \Xi(z) = \sum_{s=0}^{\infty} \xi_s z^s, \qquad \Psi_Z(z) = \sum_{s=0}^{\infty} \Psi_{Z,s} z^s.$$

The convolution system (27), $\sum_{s=0}^{S} A_s \xi_{k+s} = \Psi_{Z,k}$ for all $k \geq 0$, is equivalent to

$$A(z)\,\Xi(z) = \Psi_Z(z). \tag{28}$$

Suppose some $\xi_{s^\star}$ is invertible; reindex so that ξ_0 is invertible. Then $\Xi(z)$ admits a right inverse

$G(z) = \sum_{m\geq 0} G_m z^m$ satisfying $\Xi(z)G(z) = I$, with coefficients determined recursively by

$$G_0 = \xi_0^{-1}, \qquad G_m = -\xi_0^{-1} \sum_{\ell=1}^{m} \xi_\ell G_{m-\ell}.$$

Multiplying (28) by $G(z)$ gives $A(z) = \Psi_Z(z)G(z)$. Matching coefficients of z^s yields the unique solution

$$A_s = \sum_{m=0}^{s} \Psi_{Z,m} G_{s-m}, \qquad s = 0, \ldots, S.$$

Because $\{G_m\}$ depend only on $\{\xi_\ell\}$ and are uniquely determined when ξ_0 is invertible, the matrices $A_0, \ldots, A_S$ are uniquely identified. □

5.4 Intuition and special cases

The system (27) nests the standard cases. If $S = 0$ (static demand), then $\xi_0 = A_0^{-1}$ and $A_0 = \xi_0^{-1}$. If $S = 1$ (myopic intertemporal demand) and the flow follows a VMA(1) with $\Psi_{Z,k} = 0$ for $k > 1$, then

$$A_0\,\xi_1 \;=\; \Psi_{Z,1}, \qquad A_0\,\xi_0 + A_1\,\xi_1 \;=\; \Psi_{Z,0},$$

so that $A_0 = \Psi_{Z,1}\xi_1^{-1}$ and $A_1 = [\,\Psi_{Z,0} - A_0\xi_0\,]\xi_1^{-1}$. With longer memory in $\{\Psi_{Z,k}\}$ and higher S, the same recursion applies, moving backward through $\xi_q, \xi_{q-1}, \ldots$ to solve sequentially for $A_0, A_1, \ldots, A_S$. The economic content is that the persistence of flows produces a structured pattern of price responses across horizons; matching that pattern pins down the intertemporal substitution embodied in $\{A_s\}$.

5.5 Practical implementation and connection to instrumented elasticities

In practice, $\{\xi_s\}$ are estimated as reduced-form impulse responses of prices to the instrumented flow shock, while $\{\Psi_{Z,s}\}$ are estimated from the time–series dynamics of the instrument itself. Equation (27) then delivers $\{A_s\}$ by linear algebra. This generalizes the benchmark observation that η^{DIE} embeds a persistence wedge relative to $-\partial Q^{(p,d)}/\partial P_t^\top$: in the dynamic setting, the full stack $\{\xi_s\}$ encodes not only the level but also the horizon-by-horizon shape induced by pass-through, and combining this with $\{\Psi_{Z,s}\}$ strips out the instrument's dynamics to recover the structural demand system.

The identification relies on having an instrument that is orthogonal to preference shocks and sufficiently persistent to generate informative cross-horizon variation in $\{\xi_s\}$. When these conditions hold, the conclusion is strong: instrumented, reduced-form price responses do not merely quantify price pressure; together with the instrument's dynamics, they reveal the underlying intertemporal demand matrices $A_0,\ldots,A_S$.

6 Conclusion

This paper develops a unified framework for estimating, interpreting, and applying demand elasticities in dynamic asset markets. We begin from the simple observation that, because prices are forward-looking, there is no single "true" elasticity linking prices to demand. What an empirical design recovers depends on which expectations—about dividends, discount rates, or both—are allowed to adjust when prices move. We formalize this dependence by introducing a taxonomy of elasticities—static, immediate, and dynamic—that clarifies the relationship between structural objects, equilibrium price multipliers, and empirical estimates.

Using a tractable linear model, we show that the dynamic elasticity governs the equilibrium mapping from persistent flow shocks to prices and equals the inverse of the price multiplier. The model delivers sharp comparative statics: price multipliers rise with risk, persistence, systematic exposure, and surprise. These predictions hold empirically across two independent sources of non–cash-flow demand variation—flow-induced trading (FIT) and order flow imbalance (OFI)—demonstrating that the elasticity–multiplier link is robust and quantitatively relevant.

We then show how reduced-form price responses can be inverted to recover the underlying intertemporal demand system. A persistent instrument generates a sequence of price impulse responses that encode how investors substitute intertemporally across expected returns. Combining these responses with the measured dynamics of the instrument identifies the full set of structural demand matrices governing both contemporaneous and forward-looking portfolio behavior. This recovery result generalizes the familiar insight that the price multiplier is the inverse of the dynamic elasticity: it shows how, in principle, all components of intertemporal demand can be obtained from reduced-form data.

Taken together, our results provide a coherent foundation for interpreting empirical demand elasticities and for using them in counterfactual analysis. Dynamic settings do not undermine identification—they simply determine which elasticity is being measured. The dynamic instrumented elasticity is the correct equilibrium object for persistent shocks and the appropriate basis for counterfactual inference. More broadly, the framework developed here bridges empirical demand estimation and theoretical asset-pricing models, providing a consistent language for studying how investor flows shape prices over time.

References

Becker, Gary S, 1965, A theory of the allocation of time, *The Economic Journal* 75, 493–517.

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

Binsbergen, Jules H van, Benjamin David, and Christian C Opp, 2025, How (not) to identify demand elasticities in dynamic asset markets, Working paper.

Chaudhry, Aditya, 2025, The impact of prices on analyst cash flow expectations: Reconciling subjective beliefs data with rational discount rate variation, *Journal of Financial Economics* 171, 104095.

Chen, Andrew Y, and Tom Zimmermann, 2021, Open source cross-sectional asset pricing, *Critical Finance Review, Forthcoming* .

Chen, Tianqi, and Carlos Guestrin, 2016, Xgboost: A scalable tree boosting system, in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Cochrane, John H, 2008, The dog that did not bark: A defense of return predictability, *The Review of Financial Studies* 21, 1533–1575.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Davis, Carter, Samuli Knüpfer, Jens Soerlie Kvaerner, Bahar Sen Dogan, and Petra Vokata, 2024, Do households matter for asset prices?, Working paper.

Frazzini, Andrea, and Owen A Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299–322.

Gabaix, Xavier, and Ralph SJ Koijen, 2022, In search of the origins of financial fluctuations: The inelastic markets hypothesis, Working paper.

Haddad, Valentin, Zhiguo He, Paul Huebner, Péter Kondor, and Erik Loualiche, 2025a, Causal inference for asset pricing, Working paper.

Haddad, Valentin, Paul Huebner, and Erik Loualiche, 2025b, How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing, *American Economic Review* 115, 975–1018.

Kelly, Bryan T, Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.

Koijen, RS, Robert J Richmond, Motohiro Yogo, et al., 2019, Which investors matter for global equity valuations and expected returns, *Available at SSRN* .

Lancaster, Kelvin J, 1966, A new approach to consumer theory, *Journal of Political Economy* 74, 132–157.

Lee, Charles MC, and Mark J Ready, 1991, Inferring trade direction from intraday data, *The Journal of Finance* 46, 733–746.

Li, Jiacui, 2022, What drives the size and value factors?, *Review of Asset Pricing Studies* 12, 845–885.

Li, Jiacui, and Zihan Lin, 2023, Prices are less elastic at more aggregate levels, Working paper.

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

Tamoni, Andrea, Stanislav Sokolinski, and Yizhang Li, 2024, Which investors drive anomaly returns and how?, Working paper.

Van der Beck, Philippe, 2021, Flow-driven ESG returns, Technical report, Swiss Finance Institute.

Van der Beck, Philippe, 2025, Short versus long-run demand elasticities in asset pricing, Working paper.

Appendix

A Equivalent representations

In this section, we describe how to go from one representation of the demand system to another. We start with the representation in terms of price and dividends:

$$q_t = q^{(p,d)}(p_t, \mathbb{E}_t[p_{t+1}], \mathbb{E}_t[d_{t+1}], \mathbb{E}_t[p_{t+2}], \mathbb{E}_t[d_{t+2}], \ldots; f_t).$$

From the Campbell-Shiller approximation, we obtain

$$r_{t+1} = \kappa + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t.$$

To simplify the notation, we set $\kappa = 0$. Results extend trivially to the case where $\kappa \neq 0$. We will also omit the demand shifter f_t for brevity.

Equivalence between (p,d) **and** (p,r)**.** Solving for expected dividends at horizon h we obtain:

$$\mathbb{E}_t[d_{t+h}] = \frac{\mathbb{E}_t[r_{t+h}] + \mathbb{E}_t[p_{t+h-1}] - \rho\mathbb{E}_t[p_{t+h}]}{1-\rho}.$$

We can use the expression above to eliminate expected dividends from the demand system:

$$\begin{aligned} q_t &= q^{(p,r)}(p_t, \mathbb{E}_t[p_{t+1}], \mathbb{E}_t[r_{t+1}], \mathbb{E}_t[p_{t+2}], \mathbb{E}_t[r_{t+2}], \ldots) \\ &= q^{(p,d)}(p_t, \mathbb{E}_t[p_{t+1}], \frac{\mathbb{E}_t[r_{t+1}] + p_t - \rho\mathbb{E}_t[p_{t+1}]}{1-\rho}, \mathbb{E}_t[p_{t+2}], \frac{\mathbb{E}_t[r_{t+2}] + \mathbb{E}_t[p_{t+1}] - \rho\mathbb{E}_t[p_{t+2}]}{1-\rho}, \ldots), \end{aligned}$$

where we omitted the constant κ for brevity. This implies that the structural demand elasticities of one representation are connected to the structural demand elasticities of another representation:

$$\frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[r_{t+h}]} = \frac{1}{1-\rho}\frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+h}]}, \qquad \frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[p_{t+h}]} = \frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[p_{t+h}]} - \frac{\rho}{1-\rho}\frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+h}]} + \frac{1}{1-\rho}\frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+h+1}]},$$

for $h \geq 1$, and

$$\frac{\partial q_t^{(p,r)}}{\partial p_t} = \frac{\partial q_t^{(p,d)}}{\partial p_t} + \frac{1}{1-\rho}\frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+1}]}.$$

We can perform the reverse operation, that is, recover the (p,d) representation from the (p,r) representation. We start again from the Campbell-Shiller approximation:

$$\mathbb{E}_t[r_{t+h}] = \rho\mathbb{E}_t[p_{t+h}] + (1-\rho)\mathbb{E}_t[d_{t+h}] - \mathbb{E}_t[p_{t+h-1}].$$

We can then use the expression above to eliminate expected returns from the demand system in (p,r):

$$\begin{aligned} q_t &= q^{(p,d)}(p_t, \mathbb{E}_t[p_{t+1}], \mathbb{E}_t[d_{t+1}], \mathbb{E}_t[p_{t+2}], \mathbb{E}_t[d_{t+2}], \ldots) \\ &= q^{(p,r)}(p_t, \mathbb{E}_t[p_{t+1}], \rho\mathbb{E}_t[p_{t+1}] + (1-\rho)\mathbb{E}_t[d_{t+1}] - p_t, \mathbb{E}_t[p_{t+2}], \rho\mathbb{E}_t[p_{t+2}] + (1-\rho)\mathbb{E}_t[d_{t+2}] - \mathbb{E}_t[p_{t+1}], \ldots). \end{aligned}$$

The structural demand elasticities of the (p,d) representation are connected to the structural demand elasticities of the (p,r) representation:

$$\frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+h}]} = (1-\rho)\frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[r_{t+h}]}, \qquad \frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[p_{t+h}]} = \frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[p_{t+h}]} + \rho\frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[r_{t+h}]} - \frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[r_{t+h+1}]},$$

for $h \geq 1$, and

$$\frac{\partial q_t^{(p,d)}}{\partial p_t} = \frac{\partial q_t^{(p,r)}}{\partial p_t} - \frac{\partial q_t^{(p,r)}}{\partial \mathbb{E}_t[r_{t+1}]}.$$

Equivalence between (p,d) and (r,d). We now show that the (p,d) representation is equivalent to the (r,d) representation. We start again from the Campbell-Shiller approximation, but this time we solve the equation forward and use a no-bubble condition:

$$p_t = \sum_{j=1}^{\infty} \rho^{j-1}\left[(1-\rho)\mathbb{E}_t[d_{t+j}] - \mathbb{E}_t[r_{t+j}]\right].$$

Evaluating at horizon h and taking expectations, we obtain:

$$\mathbb{E}_t[p_{t+h}] = \sum_{j=h+1}^{\infty} \rho^{j-(h+1)}\left[(1-\rho)\mathbb{E}_t[d_{t+j}] - \mathbb{E}_t[r_{t+j}]\right] \equiv g_{t,h}(\mathbb{E}_t[r_{t+h+1}], \mathbb{E}_t[d_{t+h+1}], \ldots).$$

We can then use the expression above to eliminate expected returns from the demand system in (p,d):

$$\begin{aligned} q_t &= q^{(r,d)}(\mathbb{E}_t[r_{t+1}], \mathbb{E}_t[d_{t+1}], \mathbb{E}_t[r_{t+2}], \mathbb{E}_t[d_{t+2}], \ldots) \\ &= q^{(p,d)}(\underbrace{g_{t,0}(\mathbb{E}_t[r_{t+1}], \mathbb{E}_t[d_{t+1}], \ldots)}_{p_t}, \underbrace{g_{t,1}(\mathbb{E}_t[r_{t+2}], \mathbb{E}_t[d_{t+2}], \ldots)}_{\mathbb{E}_t[p_{t+1}]}, \mathbb{E}_t[d_{t+1}], \ldots). \end{aligned}$$

We can then use the expression above to recover the structural demand elasticities of the (r,d) representation from the structural demand elasticities of the (p,d) representation:

$$\frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[r_{t+h}]} = \sum_{j=0}^{h-1} \frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[p_{t+j}]} \frac{\partial g_{t,j}}{\partial \mathbb{E}_t[r_{t+h}]}, \qquad \frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[d_{t+h}]} = \sum_{j=0}^{h-1} \frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[p_{t+j}]} \frac{\partial g_{t,j}}{\partial \mathbb{E}_t[d_{t+h}]} + \frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+h}]}.$$

for $h \geq 1$.

We can perform the reverse operation, that is, recover the (p,d) representation from the (r,d) representation. We start again from the Campbell-Shiller approximation:

$$\mathbb{E}_t[r_{t+h}] = \rho \mathbb{E}_t[p_{t+h}] + (1-\rho)\mathbb{E}_t[d_{t+h}] - \mathbb{E}_t[p_{t+h-1}].$$

We can then use the expression above to eliminate expected returns from the demand system in (r,d):

$$\begin{aligned} q_t &= q^{(p,d)}(p_t, \mathbb{E}_t[p_{t+1}], \mathbb{E}_t[d_{t+1}], \mathbb{E}_t[p_{t+2}], \mathbb{E}_t[d_{t+2}], \ldots) \\ &= q^{(r,d)}(\rho \mathbb{E}_t[p_{t+1}] + (1-\rho)\mathbb{E}_t[d_{t+1}] - p_t, \mathbb{E}_t[d_{t+1}], \rho \mathbb{E}_t[p_{t+2}] + (1-\rho)\mathbb{E}_t[d_{t+2}] - \mathbb{E}_t[p_{t+1}], \mathbb{E}_t[d_{t+2}], \ldots). \end{aligned}$$

We can then use the expression above to recover the structural demand elasticities of the (p,d) representation from the structural demand elasticities of the (r,d) representation:

$$\frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[p_{t+h}]} = \rho \frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[r_{t+h}]} - \frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[r_{t+h+1}]}, \qquad \frac{\partial q_t^{(p,d)}}{\partial \mathbb{E}_t[d_{t+h}]} = (1-\rho) \frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[r_{t+h}]} + \frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[d_{t+h}]}.$$

for $h \geq 1$, and

$$\frac{\partial q_t^{(p,d)}}{\partial p_t} = -\frac{\partial q_t^{(r,d)}}{\partial \mathbb{E}_t[r_{t+1}]}.$$

Conclusion. We have seen that the (p,d) representation is equivalent to the (p,r) representation and the (r,d) representation. Hence, the (p,r) and (r,d) representations are also equivalent. This implies that our results on recovering the structural demand elasticities from the data are valid for all three representations.