Why is Stock-Level Demand Inelastic? A Portfolio Choice Approach*

Carter Davis[†] Mahyar Kargar[‡] Jiacui Li[§]

October 25, 2024

Abstract

Classical asset pricing models predict that optimizing investors exhibit extremely high demand elasticities, while empirical estimates are significantly lower—by three orders of magnitude. To reconcile this disparity, we introduce a novel decomposition of investor demand elasticity into two key components: "price pass-through," which captures how price movements forecast returns, and "unspanned returns," reflecting a stock's lack of perfect substitutes. In a factor model framework, we show that unspanned returns become significant when models include "weak factors." Classical models overestimate demand elasticity by assuming both very low unspanned returns and high price pass-throughs, assumptions that are inconsistent with empirical evidence.

Keywords: Demand elasticity, price pass-through, spanning, weak factors, demand-based asset pricing

JEL Classification: G11, G12, G14.

*We thank Yu An, Aditya Chaudhry, Mike Chernov, Winston Dou, Zhiyu Fu, Xavier Gabaix, Andrei Gonçalves, Valentin Haddad, Leyla Han, Christian Heyerdahl Larsen, Da Huang, Tim Johnson, Ralph Koijen, Serhiy Kozak, Daniel Neuhann, Stavros Panageas, Lubos Pastor, Markus Pelger, Nick Roussanov, Dejanir Silva, and conference and seminar participants at AFA 2024, MFA 2024, NBER New Developments in Long-Term Asset Management Spring 2024 Meeting, University of Connecticut Finance Conference, University of Maryland Young Scholars Conference, 2022 Wabash River Finance Conference, Indiana University, University of Rochester, and University of Utah for valuable feedback. We thank Andrea Tamoni and Michael Weber for sharing data. An earlier version of this paper was circulated under the title "Why is Asset Demand Inelastic?". All errors are our own responsibility.

[†]Kelley School of Business, Indiana University; ckd1@iu.edu

[‡]Gies College of Business, University of Illinois Urbana-Champaign; kargar@illinois.edu

[§]David Eccles School of Business, University of Utah; jiacui.li@eccles.utah.edu

1 Introduction

In classical asset pricing models, stock-level investor demand curves are nearly flat, implying virtually no price impacts from flows and supply shocks (e.g., Gabaix and Koijen, 2022). In stark contrast, empirical estimates indicate that flows can create sizeable price impacts and that investor demand is inelastic. This gap is large: theoretical models often predict demand elasticity in the thousands, while empirical estimates are lower by three orders of magnitude.[1] Why do investors exhibit low demand elasticities in practice? One might hypothesize that this gap arises from severe portfolio frictions or behavioral deviations from rational portfolio choice. For example, prior research has shown that investors face benchmarking incentives, leverage constraints, and transaction costs, and they also exhibit behavioral deviations from optimal portfolio choice, all of which can dampen demand elasticities (e.g., Gromb and Vayanos, 2010; Basak and Pavlova, 2013; Haddad, Huebner, and Loualiche, 2022).

In this paper, we show that, without relying on the aforementioned frictions, the bulk of the demand elasticity gap can be explained by two factors: (1) stocks are poor substitutes for each other, and (2) the limited effect of price changes on next-period returns. We analytically study investors who form optimal mean-variance (MV) portfolios and present a novel decomposition of their demand elasticity into these two components. We find that classical models predict extremely high demand elasticities due to unrealistic assumptions about the magnitude of these two factors. When we use empirically estimated values, the predicted demand elasticity decreases dramatically from approximately 7,000 to about 7, much closer to the empirical estimates around one.[2]

We begin by analytically showing that the stock-level MV investor demand elasticity decomposes

[1]To appreciate the difference in magnitude, note that for a 1% change in share supply, classical models predict a price movement of approximately 0.02 basis points (1%/5,000), which is very close to zero, whereas empirical estimates imply a price impact of around 1%.

[2]For empirical estimates of demand elasticities, see Shleifer (1986), Lou (2012), Chang, Hong, and Liskovich (2015), Koijen and Yogo (2019), Haddad et al. (2022), and Pavlova and Sikorskaya (2023), among others.

into two components. For any stock i, the elasticity of an investor's share demand (Q_i) to price (P_i) is approximately equal to one plus the product of two terms:

$$-\frac{\partial \log(Q_i)}{\partial \log(P_i)} \approx 1 + \underbrace{\left(-\frac{\partial \mu_i}{\partial \log(P_i)}\right)}_{\text{price pass-through}} \times \underbrace{\frac{\partial \log(w_i)}{\partial \mu_i}}_{=1/\mu_{i,\text{unspanned}}}, \tag{1}$$

where the first term, "price pass-through", measures the extent to which price movements unrelated to cash flows affect next-period expected returns (μ_i). The second term captures how the investor's log portfolio weight responds to changes in next-period expected return and equals the reciprocal of the "unspanned return" of stock i—defined as the residual expected return of the stock not explained by a linear projection onto other stocks. Unspanned return measures the degree to which stock i is *substitutable*. If it has near-perfect substitutes, its unspanned return will be close to zero, as assumed in classical asset pricing models like the CAPM. Otherwise, its unspanned return can be non-negligible. Importantly, both terms in Equation (1), price pass-through and unspanned return, depend only on the properties of stock returns and can be estimated from data.

Using our decomposition, we find that classical models predict extremely high demand elasticities because they assume high price pass-through and low unspanned returns. Are these assumptions supported by the data? To investigate, we estimate the two components using monthly U.S. stock return data. First, we estimate price pass-through using Fama-MacBeth regressions to predict next-month stock returns using the "price wedge" in van Binsbergen, Boons, Opp, and Tamoni (2023), which captures price variations unrelated to cash flows. We estimate price pass-through to be approximately 0.014 at the monthly horizon—meaning that each 1% drop in cash flow-unrelated prices leads to a 1.4 basis points increase in expected returns the following month. For robustness, we also consider alternative instruments aimed at isolating non-fundamental price variation, and find similar or lower estimates. Incorporating the estimated pass-through of 0.014

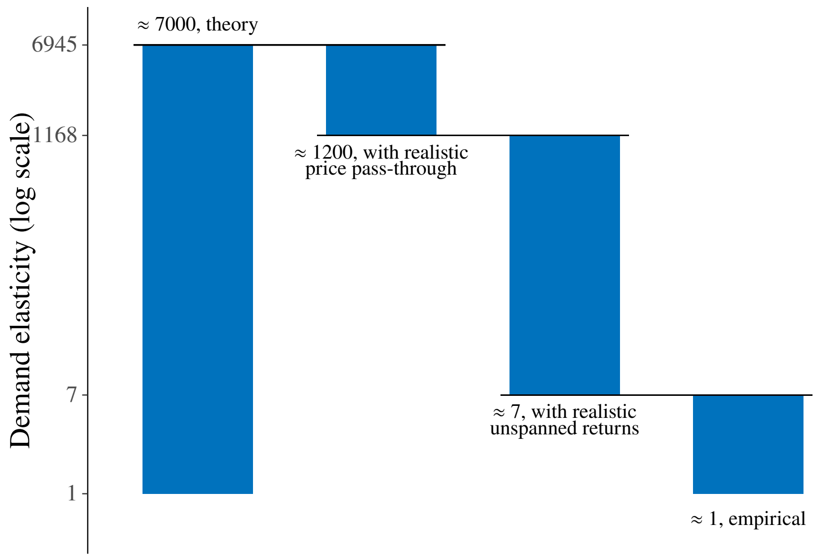

Figure 1. Explaining the gap between theoretical and empirical demand elasticities.
This figure illustrates our explanation of the gap between high demand elasticity values in classical asset pricing models and low values from empirical estimates. The leftmost bar illustrates predictions from classical models with high price pass-through and low unspanned return (e.g., Gabaix and Koijen, 2022; Petajisto, 2009, ≈ 7,000). Taking into account empirically estimated price pass-through reduces elasticity by a factor of 6 to approximately 1,200, as shown by the second bar from the left. Taking into account empirically estimated unspanned returns, we arrive at elasticity estimates of around 7, further reducing elasticity by a factor of approximately 170, as shown by the third bar from the left. Therefore, taking into account empirical estimates of price processes can explain the gap from 7,000 to 7. We hypothesize that the remaining gap from 7 to 1 (empirical estimates) may be explained by additional considerations such as transaction costs. The y-axis is in log scale.

reduces the demand elasticity prediction from 7,000 to approximately 1,200, as illustrated in the second bar of Figure 1.

We then estimate unspanned returns, which requires models of expected returns and covariance matrices for stocks. We estimate stock expected returns using Fama-MacBeth regressions with commonly used stock characteristics, and estimate covariance matrices using rolling one-year windows of daily stock returns with Ledoit and Wolf (2004) shrinkage to ensure matrix invertibility. Our results indicate that while about two-thirds of individual stock returns are spanned by linear combinations of other stocks, approximately one third remains unspanned. For an average stock with positive MV portfolio weights, we estimate its monthly unspanned return to be 0.23%. This contrasts sharply with classical models, which typically assume monthly unspanned returns around 0.001% (Gabaix and Koijen, 2022). When we incorporate empirical estimates, the predicted demand elasticity decreases sharply from 1,200 to approximately 7, as shown in Figure 1.

Thus, we find that the majority of the demand elasticity gap—from 7,000 to 7—can be explained by incorporating empirical features of stock returns. What about the remaining gap from 7 to 1? One plausible explanation is the consideration of transaction costs. However, there are many other possibilities, and we do not take a definitive stance on the exact mechanism.

For robustness, we also examine additional factors that may affect investor demand elasticity. Our two-part decomposition in Equation (1) focuses on the effects driven by changes in expected returns, but one might naturally question the impact of changes in return volatility or correlations. We empirically show that these effects tend to be small. Additionally, we consider alternative investor preferences, such as constant relative risk aversion and Epstein-Zin, which account for wealth effects and intertemporal hedging. We find that their differences from the mean-variance model are also quantitatively small for the purpose of studying demand elasticities. Finally, while demand elasticities can be mechanically high for stocks with very small portfolio weights due to the magnification effect of small denominators, when we consider more economically relevant quantities such as portfolio holdings-weighted averages, our finding of low demand elasticity remains unchanged.

Our paper finds that stocks exhibit low price pass-throughs and high unspanned returns. How can we interpret these empirical facts? The former simply implies that variations in expected return are often persistent, a common feature in many dynamic asset pricing models. The latter may seem less familiar to researchers used to factor model frameworks. To clarify the connection, we show that stocks with high unspanned returns can be understood as reflecting the presence of "weak factors", as described by Lettau and Pelger (2020). In this context, "weak factors" refer to factors that explain cross-section of expected returns but not much of the time-series variation of returns. While early factor models, such as the CAPM, assume the absence of weak factors, subsequent research found that more (and weaker) factors are necessary to describe the cross-section. Recent asset pricing models have incorporated a large number of stock characteristics (e.g., Kelly, Pruitt, and Su, 2019; Lettau and Pelger, 2020; Chen, Roussanov, and Wang, 2023). We show that these models can accommodate weak factors, and when the number of factors is sufficiently large, they invariably imply high unspanned returns and low demand elasticities.

In interpreting our results, two clarification are in order. First, we do not consider our estimate of a demand elasticity of seven to be extremely precise. For instance, investors may use different models of expected returns and covariance in practice, which could lead them to identify different unspanned returns. However, our findings remain within the same order of magnitude across various estimation techniques. Regardless of the specific estimation method, we find that using empirically estimated properties of stock returns can explain the majority of the gap between classical theory predictions and empirical estimates of demand elasticity.

Second, why do we focus on frictionless MV investors, even though many real-life investors face constraints and frictions? It is not because we believe those considerations are unimportant. Rather, we aim to ensure that our results are not driven by those additional factors, all of which can further reduce demand elasticities. We find that, once we account for empirically estimated return processes, even MV investors exhibit relatively inelastic demand. Thus, the demand of investors with constraints and frictions is likely even more inelastic. In other words, our paper provides a simple frictionless benchmark for asset-level demand elasticities, similar to how Modigliani and Miller (1958) offers a benchmark for frictionless capital structure choice. Related to this point, we clarify that our paper focuses on how optimal investor demand *theoretically* responds to price movements. While we empirically estimate statistical properties of stock returns, we do not empirically estimate demand elasticity and therefore do not need or propose a new price instrument.

This paper offers an explanation for the majority of the gap between the extremely high predictions of demand elasticity in classical models and the low empirical estimates. Specifically, we show that MV investor demand elasticity depends on two components—price pass-through and unspanned return—and that incorporating empirical estimates can lower theoretical demand elasticity predictions from 7,000 to around 7. Of the two components, incorporating empirically estimated unspanned returns have the greatest impact. Empirically, stocks are not perfect substitutes, which results in significantly lower investor demand elasticities.

There is an extensive literature estimating demand elasticities in stocks using the price impacts of demand or supply shocks: index exclusion (Shleifer, 1986; Chang et al., 2015; Pavlova and Sikorskaya, 2023), dividend payments (Schmickler, 2020), mutual fund flows (Lou, 2012), and transaction costs of realized trades (Frazzini, Israel, and Moskowitz, 2018; Bouchaud, Bonart, Donier, and Gould, 2018). More recently, structural approaches have been employed to estimate demand elasticities using asset demand systems (Koijen and Yogo, 2019; van der Beck, 2022; Haddad et al., 2022). The findings consistently show that stock-level demand is far less elastic than what classical theories predict. In this paper, we take the empirical behavior of stock returns as given and show that optimizing investors would naturally exhibit inelastic demand.

Our finding that most stock price movements exhibit low pass-through to future expected returns is consistent with existing evidence from the cross-section of stock returns. Stock returns typically exhibit reversals within a month (Jegadeesh, 1990), momentum over quarterly to annual frequency (Jegadeesh and Titman, 1993), and reversals over multiple years (De Bondt and Thaler, 1985). These effects are far less than one-for-one and are consistent with our estimates. The innovation of our paper lies not in estimating price pass-throughs, but in clarifying that weak price pass-throughs lead to low demand elasticities.

We show that demand elasticity should be low because stocks are not perfect substitutes, a topic that has received attention since the early days of capital markets theory. Scholes (1972) argues that stocks should be highly substitutable and not "unique works of art." However, our findings suggest that stocks are also not as substitutable as classical theory assumes. Early factor models, such as the CAPM or the Fama-French three factor model, assume that only systematic risk factors, which explain a large fraction of common return variation, are priced. However, subsequent research has discovered an increasing number of factors, and more recent factor models are designed to flexibly incorporate "weak factors" (e.g., Kelly et al., 2019; Lettau and Pelger, 2020; Chen et al., 2023). Researchers find that stock returns are poorly spanned by systematic risk factors (e.g., Lopez-Lira and Roussanov, 2023; Dello-Preite, Uppal, Zaffaroni, and Zviadadze, 2024) or a range of factor models (e.g., Baba Yara, Boyer, and Davis, 2021). We contribute to the literature by showing that weak factors lead to high unspanned returns, which, in turn, explain low demand elasticities.

The remainder of the paper is as follows. In Section 2, we show that MV investor demand elasticity depends on two components, price pass-through and unspanned return. In Section 3, we show that using empirical estimates of these two components lowers theoretical predictions of demand elasticities by three orders of magnitude. Section 4 discusses additional influences on demand elasticities. Section 5 shows that high unspanned returns in stocks can be understood as the existence of weak factors. Section 6 concludes.

2 Determinants of Demand Elasticities in Portfolio Choice

In this section, we show that asset-level demand elasticities in standard portfolio choice can be decomposed into two components: price pass-through and unspanned return. The former measures how price changes "pass-through" to expected returns, while the latter depends on the extent to which assets are substitutable in the cross-section. We then apply our decomposition framework to explain why classical models predict very high demand elasticities.

Why do we adopt a portfolio choice approach rather than use preference-based logistic style models, which are commonly used in the industrial organization literature? This choice is motivated by the simple insight that demand for assets is fundamentally different from demand for consumption goods, such as bananas. A consumer's demand elasticity for bananas is *directly* determined by their primitive preference for bananas. In contrast, an investor only has *indirect* preference over assets: they do not care about the assets themselves, but with the *returns* they generate, which ultimately affect their consumption or wealth. Thus, how an investor adjusts their portfolio in response to price changes primarily depends on their perception of how returns have changed.

Our subsequent results focus on the portfolio choice of mean-variance (MV) investors. For MV investors, the only factor that dampens demand elasticity is risk aversion. Therefore, we do not consider any portfolio constraints or frictions, such as benchmarking considerations and indexing, all of which tend to further reduce demand elasticities (e.g., Pavlova and Sikorskaya, 2023; Haddad et al., 2022). We make this modeling choice not because of realism, but to ensure that our results are not driven by additional frictions.

Some researchers use alternative models of investor preferences such as constant relative risk aversion (CRRA) or Epstein-Zin. These models differ from MV by incorporating intertemporal hedging considerations. Later in Section 4.2, we find that the impact of intertemporal hedging on demand elasticity is quantitatively minor. Therefore, for analytical simplicity, we focus on MV in this section.

2.1 A two-part decomposition of demand elasticity

Definition. Consider an investor who holds $Q_{i,t}$ shares of asset i at time t. The investor's demand elasticity $\eta_{i,t}$ for asset i is defined as:

$$\eta_{i,t} \equiv -\frac{\partial \log(Q_{i,t})}{\partial \log(P_{i,t})}, \tag{2}$$

where $P_{i,t}$ represents the per-share asset price, and the price change is driven by non-cash flow-related reasons. For example, an elasticity of 4 means that a 1% drop in price, unrelated to cash flows, leads to a 4% increase in the demand for shares currently held. It is important to emphasize that this definition only applies to price movements not driven by cash flows. In most models, investor demand does not react to price changes purely driven by cash flows.

To express demand elasticity in terms of portfolio weight, $w_{i,t}$, we can write $Q_{i,t} = A_t w_{i,t}/P_{i,t}$, where A_t denotes assets under management (AUM) or the investor's wealth. Substituting this into Equation (2) and assuming that A_t is exogenous,[3] we obtain:

$$\eta_{i,t} = -\frac{\partial \log(A_t w_{i,t}/P_{i,t})}{\partial \log(P_{i,t})} = 1 - \frac{\partial \log(w_{i,t})}{\partial \log(P_{i,t})}. \tag{3}$$

[3]This means that we ignore wealth effects and assume, as in Koijen and Yogo (2019), that $\frac{\partial \log(A_t)}{\partial \log(P_{i,t})} = 0$. Appendix A.2 shows that wealth effects are small for reasonably diversified portfolios.

Decomposition. We now apply the chain rule to Equation (3) to derive our decomposition based on the effect of changes in expected return $\mu_{i,t}$:

$$\eta_{i,t} \approx 1 + \underbrace{\frac{\partial \log(w_{i,t})}{\partial \mu_{i,t}}}_{\text{weight responsiveness}} \times \underbrace{\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right)}_{\text{price pass-through}}. \tag{4}$$

The first term, $\frac{\partial \log(w_{i,t})}{\partial \mu_{i,t}}$, which we refer to as "weight responsiveness," describes how responsive the investor's portfolio weights are to expected returns. The second term, $-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}$, which we refer to as "price pass-through," measures the fraction of price movements that are "passed through" to changes in expected return. For instance, if a 1% decrease in price leads to a 0.3% increase in next-period expected return, the price pass-through would be 0.3.

The equality in Equation (4) is approximate because our decomposition only considers the demand response to changes in expected returns. Of course, price changes can also lead to changes in other factors, such as the asset's return volatility or its consumption-hedging properties, all of which may also impact demand. However, under reasonable assumptions, as shown in Section 4, these additional terms are quantitatively small. Therefore, in this section, we focus primarily on the effects driven by expected returns.

2.2 Weight responsiveness depends on unspanned returns

What determines the two terms in Equation (4)? Price pass-through clearly depends on price dynamics. If a price movement is expected to largely revert in the next period, the pass-through will be close to one. However, if the price movement is more persistent, the pass-through will be lower.

The determinants of weight responsiveness are less clear. In this section, we show that for an MV investor, weight responsiveness is inversely proportional to "unspanned return", which measures the asset's distinctiveness.

Portfolio choice problem. We consider a standard portfolio choice problem with constant absolute risk aversion (CARA) utility and multivariate normally distributed returns, which results in mean-variance (MV) portfolio choice. Suppose there are N risky assets and a risk-free asset with an exogenously given gross return $R_{f,t}$. The investor's optimization problem is:

$$\max_{w_t} \mathbb{E}_t \left[-\exp\{-\gamma A_t \left(w_t' r_{t+1} + R_{f,t} \right)\} \right],$$

where γ is the coefficient of absolute risk aversion, w_t is an N dimensional vector of portfolio weights for risky assets, and r_t is an N dimensional vector of excess returns on the risky assets. Investor wealth, A_t, is exogenously given. The first-order condition yields the familiar MV portfolio weights:

$$w_t = \frac{1}{\gamma A_t} \Sigma_t^{-1} \mu_t, \tag{5}$$

where Σ_t is the $N \times N$ covariance matrix of returns, and μ_t is the N dimensional vector of expected excess returns.

Unspanned return determines weight responsiveness. To determine asset-level weight responsiveness, we need to study how the portfolio weight in asset i, $w_{i,t}$, responds to $\mu_{i,t}$. For this, Stevens (1998) suggests that the key is isolating the component of asset i that is not spanned by other assets. Specifically, consider the regression of asset i excess return, $r_{i,t+1}$, on the $N-1$ dimensional vector of excess returns of all other assets, $r_{-i,t+1}$:

$$r_{i,t+1} = \underbrace{\mu_{i,\text{unspanned},t}}_{\text{non-replicable part}} + \underbrace{\beta'_{-i,t} r_{-i,t+1}}_{\text{replicating portfolio}} + \epsilon_{i,t+1}, \tag{6}$$

where $\beta_{-i,t}$ is the $(N-1)$ dimensional vector of conditional slope coefficients, $\epsilon_{i,t+1}$ is the mean-zero error term, and $\mu_{i,\text{unspanned},t}$ is the conditional intercept term—representing the component of asset i's return that cannot be replicated by other assets. Taking time-t conditional expectation of both sides yields

$$\mu_{i,t} = \mu_{i,\text{unspanned},t} + \underbrace{\beta'_{-i,t} \mu_{-i,t}}_{\text{spanned returns}}, \tag{7}$$

where $\mu_{i,t}$ and $\mu_{-i,t}$ are the conditional expectations of $r_{i,t+1}$ and $r_{-i,t+1}$, respectively.

Our key result, Proposition 1, shows that the weight responsiveness for asset i is equal to the reciprocal of $\mu_{i,\text{unspanned},t}$.[4]

Proposition 1. *If the investor forms MV efficient portfolios, then for all assets with positive portfolio weights* ($w_{i,t} > 0$)*, weight responsiveness is given by*

$$\frac{\partial \log(w_{i,t})}{\partial \mu_{i,t}} = \frac{1}{\mu_{i,unspanned,t}}. \tag{8}$$

Thus, the demand elasticity takes the following form:

$$\eta_{i,t} = 1 + \underbrace{\frac{1}{\mu_{i,unspanned,t}}}_{\textit{weight responsiveness}} \times \underbrace{\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right)}_{\textit{price pass-through}}. \tag{9}$$

[4]In Appendix D, we use a two-asset example to illustrate why unspanned returns—which capture the degree of substitutability between assets—matter for asset-level demand elasticity for an MV investor.

Proof. See Appendix A.1. □

Equation (9) follows directly from Equations (4) and (8). In essence, Equation (8) shows that the sensitivity of an investor's portfolio to changes in expected return is inversely proportional to the asset's unspanned return. Therefore, investor demand is more inelastic for assets that are more distinctive and have larger unspanned returns.

Where does Equation (8) come from? In the proof of Proposition 1, we show that MV demand for asset i can be written as:

$$w_{i,t} = \frac{1}{\gamma A_t}\left(\frac{\mu_{i,\text{unspanned},t}}{\sigma^2_{i,\text{unspanned},t}}\right), \tag{10}$$

where the numerator is the unspanned return of asset i, and the denominator, $\sigma^2_{i,\text{unspanned},t}$, is the conditional variance of $\epsilon_{i,t+1}$ in Equation (6), which we refer to the "unspanned variance" of asset i. This expression is similar to the result one would obtain if there were only a single risky asset in the economy, where demand is linear in the expected return divided by the return variance. The only difference is that, instead of *standalone* expected return and variance of asset i, Equation (10) involves the *unspanned* return and variance of asset i, which arises from considering the impact of other assets on portfolio choice.

If we assume that the unspanned variance, $\sigma^2_{i,\text{unspanned},t}$, remains constant, applying the logarithm and then taking the derivative with respect to $\mu_{i,t}$ yields Equation (8). Later, we relax this assumption and also consider changes in unspanned variance and betas ($\beta_{-i,t}$) in Section 4.1, finding that the difference is quantitatively small.

Intuition and clarifications. What does unspanned return, $\mu_{i,\text{unspanned},t}$, measure? We propose that it captures the distinctiveness, or non-substitutability, of asset i. The higher $\mu_{i,\text{unspanned},t}$, the less replicable the asset is by other assets, making it a more important position in the MV investor's portfolio.

As an intuitive example, consider an investor who believes that Tesla stock has high expected returns. What is the investor's demand elasticity on Tesla? Our result suggests that it depends on whether the investor views Tesla as a unique investment opportunity. If the investor believes that a significant portion of Tesla's high expected returns is specific to Tesla and cannot be achieved using other assets, her demand elasticity for Tesla will be low. Conversely, if the investor thinks Tesla's returns can be replicated by a combination of other stocks, her demand elasticity will be high.

It is important to clarify that unspanned return is not the same as asset pricing alpha in factor models such as the CAPM. In the CAPM, the explanatory variable—the market portfolio—is constructed using *all* assets, while in Equation (7), asset i itself is excluded from the explanatory variables. Therefore, even in a world where the CAPM holds, $\mu_{i,\text{unspanned},t}$ can be non-zero, albeit small, for all assets, despite the CAPM alpha being zero by definition. Section 5 further explains how to interpret unspanned returns within the traditional framework of factor models.

2.3 Why do classical models predict extremely high demand elasticities?

In this section, we offer a benchmark calibration that shows how classical assumptions yield demand elasticities of approximately 7,000. While this is a specific calibration, it reflects the key features of many classical models. Gabaix and Koijen (2022) reach a similar conclusion, showing that asset-level demand elasticities in classical pricing models are on the order of 5,000 or higher. Similarly, Petajisto (2009) finds a demand elasticity exceeding 6,000.

Our goal is to explain *why* these models predict high demand elasticities. Using our decomposition framework, we show that this is because classical models assume high price pass-throughs and, more importantly, low unspanned returns. Below, we outline the steps in our calibration.

1. **Price pass-through.** Many asset pricing models used for studying demand elasticities are

static, assuming that payoffs are realized in the next period. As a consequence, they assume that all current period price movements must revert, leading to a price pass-through of one over a period.[5] In the monthly calibration below, we follow this practice and assume that all price movements revert within one year, implying a monthly pass-through of 1/12 (e.g., Petajisto, 2009).

2. **Unspanned returns.** More importantly, classical models often implicitly assume that unspanned returns are very small. To provide a calibration, we solve for the unspanned return of an asset i:

$$\mu_{i,unspanned,t} = \left(\iota' \Sigma_t^{-1} \mu_t\right) \sigma^2_{i,\text{unspanned},t} \left(\frac{w_{i,t}}{\sum_{j=1}^N w_{j,t}}\right), \tag{11}$$

where ι is a column vector of ones. The expression above is derived by solving Equation (10) for the unspanned return and using Equation (5) to substitute out γA_t.

We use Equation (11) to calibrate for unspanned returns using values that are often assumed for U.S. stocks (e.g. Petajisto, 2009). Suppose there are $N = 1{,}000$ stocks, each with a monthly return volatility of 10% and a monthly expected excess return of 0.5%. Further, assume that all pairwise correlations are 0.3. Under these assumptions, the R^2 from regressing the return of asset i on all other assets, as in Equation (6), is approximately 30%, leading to an unspanned volatility of $\sigma_{i,t,\epsilon} = \sqrt{0.7 \times 0.1^2} \approx 8.4\%$. The average portfolio weight, $w_{i,t}/(\sum_{j=1}^N w_{j,t})$, is

[5] Even in dynamic settings, if one assumes that future prices are fixed—making the current period price effect fully temporary—then pass-through is also approximately one. To see this, consider the Campbell and Shiller (1988) decomposition:

$$\log(R_{i,t+1}) \approx \kappa_{i,t} + \delta \log(D_{i,t+1}) + (1-\delta)\log(P_{i,t+1}) - \log(P_{i,t}),$$

where $R_{i,t+1}$ is the gross return on the asset, $\kappa_{i,t}$ is the first order approximation constant, and δ is the dividend to price-plus-dividend ratio. If $P_{i,t+1}$ and $D_{i,t+1}$ are fixed, then the derivative with respect to the log-price ($\log(P_{i,t})$) must be -1, implying a pass-through is 1.

simply $1/N = 1/1{,}000$. The first scalar in Equation (11) is $\iota' \Sigma_t^{-1} \mu_t \approx 1.66$. Combining these values gives an unspanned return for this asset of 0.0012% ($\approx 1.66 \times 0.084^2 \times (1/1{,}000)$).[6]

Unspanned returns of 0.0012% indicate that stocks are highly substitutable. For comparison, the monthly average excess return of stocks is on the order of 0.5%, which is two orders of magnitude higher. Therefore, such classical calibrations assume that only a tiny fraction of expected return is unique to stock i itself, while the vast majority of its expected return can be replicated using a linear combination of other stocks (e.g., Scholes, 1972).

We are now ready to calibrate the implied demand elasticity of classical models. Substituting the calibrations of price pass-throughs and unspanned returns into our decomposition in Equation (4) yields a demand elasticity of $1 + \frac{1/12}{0.0012\%} \approx 7{,}000$. Naturally, demand elasticity depends on the model parameters, but experimenting with a range of parameter values typically produces an average elasticity across assets in the thousands. As mentioned earlier, our conclusion is consistent with Petajisto (2009) and Gabaix and Koijen (2022).

3 Empirical Estimates

In Section 2, we showed that demand elasticities for MV investors are determined by price pass-throughs and unspanned returns. Notably, both terms depend solely on the properties of stock returns. We empirically estimate these two terms in this section.

It is important to clarify that we do not claim our estimates are perfectly precise. Rather, they provide insights into *reasonable ranges* for price pass-throughs and unspanned returns. However, across various estimation approaches, the estimates differ from those assumed in classical models.

[6]For simplicity, this calibration assumes all stocks are symmetric. Appendix B.2 shows that calibrations with heterogeneous stocks yield similar results.

Variable	N	Mean	SD	5%	25%	50%	75%	95%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Market cap ($ bn)	1,633	4.58	13.96	0.30	0.51	1.08	3.02	17.65
Monthly excess return	1,633	0.72%	9.77%	-13.83%	-4.68%	0.38%	5.67%	16.24%
Lagged monthly return	1,633	1.49%	9.88%	-12.82%	-4.16%	0.86%	6.31%	17.69%
$\log(B/M)$	1,633	-0.59	0.82	-2.02	-1.06	-0.52	-0.06	0.63
Asset growth	1,633	0.13	0.19	-0.09	0.03	0.09	0.19	0.52
Dividend/book	1,633	2.91%	2.84%	0.00%	0.74%	2.11%	4.30%	9.36%
Profitability	1,633	0.24	0.22	-0.03	0.15	0.24	0.34	0.59

Table 1. Summary statistics
The sample consists of CRSP monthly U.S. stocks from 1970 to 2019. We excluded stocks smaller than the monthly NYSE 20% percentiles. The table reports the average distributions of variables across months. The first column reports the average number of stocks in each month. The next two columns report the mean and the standard deviation, and the last five columns report percentiles.

Incorporating these estimates explains a substantial portion of the discrepancy between theoretical predictions and empirical estimates of demand elasticities.

3.1 Data

We use the standard CRSP-Compustat merged dataset for returns and basic stock characteristics, with the sample spanning from 1970 through 2019. To prevent our results from being driven by microcap stocks, we exclude stocks with market capitalization lower than the 20$^{\text{th}}$ percentile of NYSE stocks each month. Quarterly and annual returns are calculated by compounding monthly returns. We obtain risk-free rate and factor returns from Ken French's website. The return-predicting stock characteristics data is from Freyberger, Neuhierl, and Weber (2020). Table 1 reports the summary statistics, with an average of 1,633 stocks included per month.

3.2 Estimates of price pass-through

As discussed in Section 2.1, price pass-through ($-\partial \mu_{i,t}/\partial \log(P_{i,t})$) measures the extent to which price variation, unrelated to cash flows, forecasts next-period expected returns. The main empirical challenge in estimating this term is isolating the price variation that is unrelated to cash flows. To address this, we use the "price wedge" measure from van Binsbergen et al. (2023), which is specifically designed for this purpose.

The price wedge measure. We briefly discuss the methodology in van Binsbergen et al. (2023) and refer readers to their paper for details. For each stock, their goal is to compute its "fundamental price", $\tilde{P}_{i,t}$, via a present-value framework:

$$\tilde{P}_{i,t} = E_t\left[\sum_{s=1}^{J} \frac{m_{t+s}}{m_t} \cdot D_{i,t+s} + \frac{m_{t+J}}{m_t} P_{i,t+J}\right],$$

where m_t is the stochastic discount factor implied by the CAPM, $D_{i,t+s}$ denotes dividend payments s periods ahead, $J = 15$ years is the terminal period, and $P_{i,t+J}$ is the terminal stock price. van Binsbergen et al. (2023) define the "price wedge" as:

$$PW_{i,t} = \log\left(\frac{P_{i,t}}{\tilde{P}_{i,t}}\right),$$

where $P_{i,t}$ is the actual price of stock i at time t. Thus, for example, a price wedge $PW_{i,t} = 5\%$ indicates that the current stock price is 5% higher than the model-implied fundamental value.

Crucial for our analysis, van Binsbergen et al. (2023) estimate price wedges using portfolios sorted on 57 cross-sectional stock characteristics known to predict returns, and then project these portfolio-level price wedges onto the stock level. As a result, the price wedges are approximately equal to the present value of future expected excess returns above that implied by the CAPM. Stocks with high (low) price wedges are, by construction, those with low (high) subsequent excess returns. In other words, price wedges are designed to isolate price movements that are unrelated to cash flows.

Estimating price pass-through. We estimate price pass-through using Fama-MacBeth regressions:

$$r_{i,t+1\to t+H} = \alpha_H - \beta_H \cdot \underbrace{\log(P_{i,t}/\tilde{P}_{i,t})}_{=PW_{i,t}} + \epsilon_{i,t+1\to t+H}, \tag{12}$$

where $r_{i,t+1\to t+H}$ is the log stock return over months $t+1$ through $t+H$. We estimate standard errors using the Newey-West procedure with H lags.

The estimated price pass-through β_H for horizons $H = 1$, 3, 6, and 12 months are reported in the first four columns of the first row of Table 2. We find positive price pass-through that increases with the horizon, although the estimate at 12-month horizon loses statistical significance due to larger standard errors. To make the estimates comparable across horizons, columns (6) through (9) report the pass-through per month (β_H/H), which turns out to be roughly constant across horizons at approximately 0.014, with the tightest 95% confidence interval ranging from 0.008 to 0.02. This estimate implies that if a stock's price is 1% lower due to cash flow-irrelevant reasons, its subsequent expected return is higher by 0.014% in the following month.[7]

Robustness to alternative price variation measures. We also consider alternative measures of cash flow-unrelated price variation. Bartram and Grinblatt (2018) measure the fundamental value-justified price $\tilde{P}_{i,t}$ using "kitchen sink" regressions that incorporate various accounting variables. We apply their methodology to compute price wedges and report the corresponding price pass-throughs in the second row of Table 2. The point estimates are all positive but statistically

[7]The price wedge in van Binsbergen et al. (2023) aggregates price variations associated with many characteristics. In Appendix B.1, we separate the characteristics into several types and estimate the price pass-throughs associated with each. The results generally confirm low price pass-through for most types of characteristics.

insignificant, and the magnitudes are smaller than those associated with the van Binsbergen et al. (2023) price wedge.[8]

We also consider two price instruments to isolate cash flow-unrelated price variation. The first one follows Koijen and Yogo (2019), who develop an instrument for the market capitalization of stocks using institutional holdings. We download their data and use their instrument in a two-step procedure. Specifically, we regress the log of stock market capitalization on the log of their instrument, and then use the predicted values as the independent variable in the Fama-MacBeth regression (12) to estimate price pass-through. The results, reported in the third row of Table 2, indicate a slightly negative price pass-through. In other words, the price variation isolated by the Koijen and Yogo (2019) instrument exhibit slight momentum rather than reversals.

The second instrument is the flow-induced trading (FIT) measure in Lou (2012). In short, FIT captures the mechanical trading responses by mutual funds as they scale up or down existing holdings in response to fund flows. The resulting FIT variable has been shown to create price effects that tend to revert over time. In the first stage, we regress the log of stock price on the most recent four quarterly values of FIT. The second stage Fama-MacBeth results, shown in the last row of Table 2, indicate a positive pass-through of slightly less than 0.01 per month, though the results are not statistically significant.

Overall, across four different specifications of cash flow-unrelated price variation, we do not find very strong price pass-throughs. The highest monthly pass-through, associated with the van Binsbergen et al. (2023) price wedge measure, is 0.014. This is much lower than typical theoretical calibrations, such as in Petajisto (2009), which assumes a monthly price pass-through of 1/12

[8]This is largely expected, as the model in Bartram and Grinblatt (2018) is built to estimate fundamental valuation and only yields return predictability as a secondary outcome, whereas the price wedges in van Binsbergen et al. (2023) are directly estimated using characteristics known to predict returns. The degree of return predictability implied in our price pass-through estimate is consistent with the original findings in Bartram and Grinblatt (2018).

Independent variable	Estimated coefficient β_H				Obs	Implied monthly price pass-through (β_H/H)			
	$H = 1$	3	6	12		$H = 1$	3	6	12
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
van Binsbergen et al. (2023)	0.014***	0.040***	0.079**	0.157	1,270,646	0.014***	0.013***	0.013**	0.013
price wedge	(0.003)	(0.014)	(0.038)	(0.102)		(0.003)	(0.005)	(0.006)	(0.009)
Bartram and Grinblatt (2018)	0.001	0.002	0.003	0.006	782,431	0.001	0.001	0.000	0.000
price wedge	(0.001)	(0.002)	(0.005)	(0.012)		(0.001)	(0.001)	(0.001)	(0.001)
Koijen and Yogo (2019)-	−0.002***	−0.004***	−0.009***	−0.021*	1,519,519	−0.002***	−0.001***	−0.002***	−0.002*
instrumented $\log(P_{i,t})$	(0.000)	(0.002)	(0.003)	(0.011)		(0.000)	(0.001)	(0.001)	(0.001)
FIT-instrumented	0.006	0.020	0.040	0.094	1,443,296	0.006	0.007	0.007	0.008
$\log(P_{i,t})$	(0.005)	(0.016)	(0.045)	(0.084)		(0.005)	(0.005)	(0.008)	(0.007)
Note:									*p<0.1; **p<0.05; ***p<0.01

Table 2. Estimating price pass-through using different measures of cash flow-unrelated price variation. We estimate price pass-through using Fama-MacBeth regressions:

$$r_{i,t+1\to t+H} = \alpha_H - \beta_H \cdot \log\left(P_{i,t}/\tilde{P}_{i,t}\right) + \epsilon_{i,t+1\to t+H},$$

where the dependent variable is the log return of stocks from months $t+1$ to $t+H$. The independent variable captures cash flow-unrelated price variation as of time t, where $\tilde{P}_{i,t}$ is a measure of fundamental valuation, and $P_{i,t}$ is the actual price. In the first two rows, the independent variables are the "price wedge" measures in van Binsbergen et al. (2023) and Bartram and Grinblatt (2018), respectively. In the last two rows, the independent variables are the log market capitalization of stocks instrumented by the Koijen and Yogo (2019) instrument and the flow-induced-trading (FIT) measure in Lou (2012), respectively. Columns (1) through (4) report the estimated regression coefficients β_H for horizons $H = 1, 3, 6$, and 12 months. Column (5) reports the number of stock-months used in each regression. Columns (6) through (9) report β_H/H, which can be interpreted as the *monthly* price pass-through estimate. Throughout, standard errors of the Fama-MacBeth forecasting coefficients are calculated using the Newey-West procedure, with the number of lags equal to the forecasting horizon H.

(annual pass-through of 1). Therefore, incorporating empirically estimated price pass-through would reduce the predicted demand elasticity by a factor of approximately $\frac{1/12}{0.014} \approx 6$.

3.3 Estimates of unspanned returns

We now estimate the second term—unspanned returns—in our decomposition of demand elasticity. Recall that, for each asset i, its unspanned return is defined as the component of its expected return that is not spanned by other assets in a linear regression. To compute this, we need to estimate both the vector of expected excess stock returns μ_t and the covariance matrix, Σ_t. In this section, we present one set of estimates for brevity, but Section 5.2 shows that our main conclusion is robust to alternative estimation approaches.

To estimate μ_t, we follow a long literature on the cross-section of stock returns (e.g., Lewellen, 2015) and predict expected returns using stock characteristics in Fama-MacBeth regressions. We estimate a full-sample Fama-MacBeth regression for simplicity, but the results are similar when using rolling window regressions. The characteristics include beta, size, book-to-market, investment, profitability, momentum, and past one-month returns. As the characteristics vary over time, so do the return forecasts.

To estimate the time-varying covariance matrix Σ_t, we follow Lopez-Lira and Roussanov (2023) by using a one-year lag of daily returns, and then slightly shrink the sample covariance matrix using the methodology in Ledoit and Wolf (2004).[9] Specifically, we use $\Sigma_t = (1-h)\hat{\Sigma}_t + h\overline{\Sigma}_t$, where $\hat{\Sigma}_t$ is the sample covariance matrix estimate, $\overline{\Sigma}_t$ is the shrinkage target, and h is the scalar shrinkage weight. The shrinkage target $\overline{\Sigma}_t$ has the average stock-specific return variance along the diagonal ($\frac{1}{N}\sum_{i=1}^{N}\hat{\Sigma}_{i,i,t}$) and the average covariance on the off-diagonal ($\frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}\hat{\Sigma}_{i,j,t}$). We use shrinkage weight $h = 0.05$ but the results are not sensitive to reasonable variations.

Implied unspanned returns. We then use our estimates of μ_t and Σ_t to compute the implied spanned and unspanned returns ($\beta'_{-i,t}\mu_{-i,t}$ and $\mu_{i,t} - \beta'_{-i,t}\mu_{-i,t}$).[10] We compute these for each stock-month and graphically illustrate the results in Figure 2. We sort the sample by each stock's total expected excess return, $\mu_{i,t}$, into 100 bins and plot the average spanned component $\beta'_{-i,t}\mu_{-i,t}$ on the y-axis. The gap between the 45-degree black dotted line and the red line labeled "Estimated"

[9]There are two reasons for shrinking the covariance matrix. First, because we have more stocks than the number of days in a year, the sample covariance matrix is not full-rank. This poses a problem, as computing unspanned returns and forming mean-variance efficient portfolios requires inverting the covariance matrix. Shrinkage ensures that the matrix is full-rank. Second, and more importantly, it is well-known that using the sample covariance matrix with a large number of assets leads to poorly conditioned portfolio choices (e.g., Ledoit and Wolf, 2004; Brandt, 2010). Shrinkage is a common regularization technique that produces implementable portfolios.

[10]The proof of Proposition 1 in Appendix A.1 shows how to calculate unspanned returns from μ_t and Σ_t.

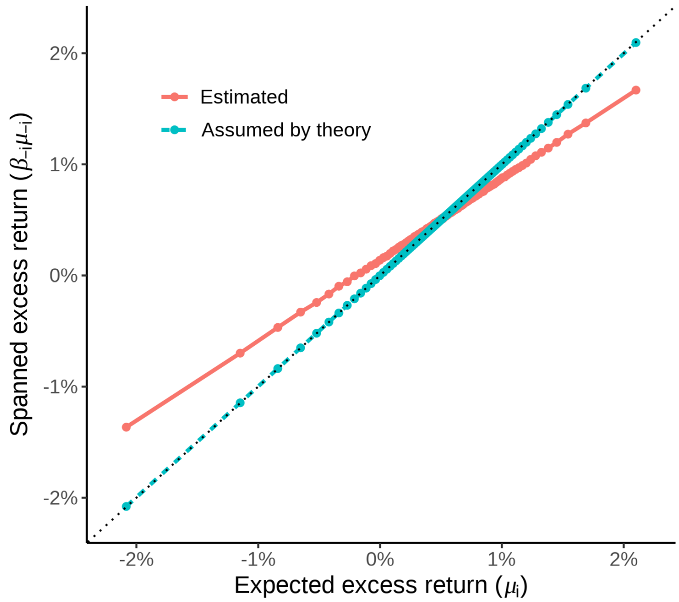

Figure 2. Decomposing spanned and unspanned returns
This figure examines the fraction of expected excess return that is spanned by other stocks. As described in Section 3.2, we estimate expected excess returns μ_t using stock characteristics-based Fama-MacBeth predictions. We estimate the covariance matrix Σ_t using daily returns in the previous 12 months, with a minor Ledoit and Wolf (2004) shrinkage to ensure positive definiteness (shrinkage parameter = 0.05). We sort the full sample into 100 bins based on each stock's estimated excess return $\mu_{i,t}$ and plot the average spanned component ($\beta'_{-i}\mu_{-i,t}$) on the vertical axis. The red line represents the spanned returns estimated from data and, for comparison, the blue line shows the unspanned returns assumed by classical theory.

represents the asset's unspanned return. When the black dotted line is above (below) the red, the unspanned return is positive (negative). For example, in the top bin, the monthly excess return is 2.10%, while the spanned component is 1.67%, resulting in an unspanned return of 2.1% − 1.67% = 0.43%.

To highlight the difference with classical models, we also plot the spanning implied by the model of Petajisto (2009) in Figure 2, shown in teal color. In that model, which is representative of classical models, approximately 99.8% of all returns are spanned, meaning the line essentially overlaps with the 45-degree line.

Overall, our empirical results indicate that while a substantial fraction of expected return is spanned, the spanning is far from complete. For stocks with positive MV portfolio weights, we calculate the average monthly unspanned return to be 0.23%.[11] In Table 3, we show that our finding is robust to using alternative Ledoit and Wolf (2004) shrinkage parameters for estimating the covariance matrix or simply using the Fama-French-Carhart (FFC) four-factor model to forecast returns. When we estimate the Fama-MacBeth regressions using 10-year rolling windows instead of the full sample, unspanned returns increase slightly. Later, Section 5.2 shows that the conclusion of non-trivial unspanned returns holds robustly across several alternative factor pricing models that capture a wide range of characteristics-implied return predictability.

In columns (5) through (8) of Table 3, we report the realized annual Sharpe ratios of the MV portfolios that trade on these return predictors. In our implementation, we scale the portfolio weights to have the same L1 norm across all periods. When using all six characteristics mentioned in Section 3.3 to forecast returns, these strategies achieve Sharpe ratios of around 1.2 to 1.4 when expected returns are estimated using the full sample, and around 1.1 to 1.3 when estimated using rolling windows, which is a more realistic approach. The Sharpe ratio declines to a range of 0.6 to 0.7 when only using the characteristics used to construct the FFC four factors. These results are comparable to, or lower than, those obtained in other academic studies that use stock characteristics to predict returns (e.g., Kelly et al., 2019; Kim, Korajczyk, and Neuhierl, 2021).

3.4 Demand elasticity implied by empirical estimates

We now show that incorporating empirical estimates of price pass-throughs and unspanned returns significantly reduce the demand elasticity prediction. This process is visually shown in the waterfall graphic in Figure 1.

[11]Since demand elasticities are computed using logs, they are only well-defined for assets with positive portfolio weights. See Appendix C.1 for a discussion on extending this to cases with negative portfolio weights.

Expected return	Monthly unspanned return				Sharpe ratio			
	All characteristics		FFC		All characteristics		FFC	
	Full sample	Rolling	Full sample	Rolling	Full sample	Rolling	Full sample	Rolling
Shrinkage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.010	0.233%	0.285%	0.227%	0.265%	1.261	1.152	0.632	0.698
0.025	0.233%	0.285%	0.227%	0.265%	1.272	1.161	0.634	0.700
0.050	0.233%	0.284%	0.227%	0.265%	1.292	1.175	0.638	0.703
0.100	0.232%	0.283%	0.226%	0.264%	1.332	1.205	0.645	0.708
0.150	0.231%	0.282%	0.225%	0.263%	1.374	1.237	0.652	0.714
0.200	0.230%	0.281%	0.225%	0.262%	1.418	1.269	0.660	0.720

Table 3. Unspanned return and Sharpe ratio
This table reports the unspanned return and annual portfolio Sharpe ratio for different mean-variance portfolio implementations. We estimate the return covariance matrix using one-year lag of daily returns and apply Ledoit and Wolf (2004) shrinkage, with results for different shrinkage parameters presented in different rows. Columns (1) through (4) report the average monthly unspanned returns of stocks with positive portfolio weights. To estimate expected returns, we use Fama-MacBeth regressions to predict stock returns using either a larger set of characteristics (beta, size, book-to-market, profitability, investment, momentum, reversal) in columns (1) through (2) or just the Fama-French-Carhart (FFC) four characteristics in columns (3) through (4). Columns (5) through (8) report the resulting portfolio-level annual Sharpe ratios. In forming the mean-variance efficient portfolios over time, we scaled the portfolio so that the L1 norm of portfolio weights is equal through all time periods.

At the left of the figure, we start from the classical calibration in Section 2.3, which yields a demand elasticity of approximately 7,000. First, we incorporate our price pass-through estimate into the decomposition (Equation 4), lowering demand elasticity to approximately 1,200 ($\approx 1 + 0.014/0.000012$), as illustrated by the second bar in Figure 1. Next, we apply our empirical estimate of unspanned returns of 0.23%, which is significantly larger than the 0.0012% assumed in classical models. This further reduces demand elasticity to approximately 7 ($\approx 1 + 0.014 / 0.0023$), as illustrated by the next drop in the waterfall chart. While using estimated pass-throughs reduces demand elasticity by a factor of approximately 6, using empirical estimates of unspanned returns decreases it by another two orders of magnitude.

Our primary specification in the paper studies MV investors who rebalance portfolios monthly. However, some investors in practice rebalance less frequently and have longer investment horizons. Appendix C.3 shows that the demand elasticity predictions do not change significantly with reasonable variations in rebalancing frequency. While longer investment horizons necessarily mean higher price pass-throughs, they are also associated with larger unspanned returns, and these two effects roughly offset each other in the demand elasticity formula of Equation (9).

Negative or small portfolio weights. Up to this point, we have followed Koijen and Yogo (2019) to only compute demand elasticities for stocks with positive portfolio weights, as demand elasticities—defined using the log of holdings—are only defined for stocks with positive holdings. However, an MV investor would naturally hold short positions. It would be desirable to also consider these negative positions.

In Appendix C.1, we use a generalized definition of demand elasticity that includes short positions. Under this generalized definition, we show that our two-component decomposition still holds if we replace unspanned returns with their absolute value in Equation (9). Since the distribution of the absolute value of unspanned returns is roughly symmetric between positive and negative holdings, as shown in Figure 2, the (generalized) demand elasticities for short positions are similar to those for long positions.

So far, we use the same average unspanned returns across stocks for simplicity, and find that the demand elasticity is low. A discerning reader may note that stocks have heterogeneous unspanned returns, and those with near-zero unspanned returns would exhibit much higher demand elasticities (Equation (9)). That is true, but since the MV investor's portfolio weight is proportional to unspanned returns (Equation (10)), the investor would hold negligible positions in those stocks. In Appendix C.2, we show that positions with very small portfolio weights have small impact on aggregate demand elasticities. When considering more economically relevant measures, such as the portfolio size-weighted average demand elasticity, we still obtain an estimate of approximately 7, even with heterogeneous unspanned returns across stocks.

Understanding low price pass-throughs and high unspanned returns. As detailed in Appendix A.3, low price pass-throughs suggest that the variation in expected returns is relatively persistent. While we do not adopt a specific stance on the microfoundations underlying this persistence, it is a natural prediction of many dynamic asset pricing models. It is important to recognize that stocks are long-lived assets, and the use of a one-period model, such as in Petajisto (2009), inevitably leads to the assumption that price pass-through equals one.

What does it mean for unspanned returns to be high? Section 5 provides additional interpretations of this finding. In the context of factor models, we show that high unspanned returns are linked to the presence of many "weak factors", defined as factors that explain subsets of stocks but not necessarily the broad cross-section (Lettau and Pelger, 2020; Lopez-Lira and Roussanov, 2023). Empirically, we also show that several modern factor pricing models that incorporate a multitude of stock characteristics (and thus exhibit weak factors) predict high unspanned returns. Therefore, our finding of high unspanned returns is not specific to the model used in this section.

4 Additional Considerations

Our analysis of demand elasticity so far has focused on how MV investor demand responds to changes in expected returns. In principle, price movements can also lead to changes in volatilities and covariances, which in turn affect demand. Additionally, long-horizon investors are concerned about hedging future consumption risk. We study these effects in Sections 4.1 and 4.2, and find that their impact is small under reasonable calibrations.

As shown by the estimates in Section 3, incorporating empirically estimated properties of stock returns reduces demand elasticity predictions to around 7. While this is significantly lower than the classical theoretical calibrations, which are in the thousands, it remains higher than most empirical estimates, which are typically around one. What accounts for the remaining gap? For completeness, Section 4.3 suggests that introducing transaction costs is one plausible explanation. Other possibilities exist, and we do not take a definitive stance on the correct mechanism.

4.1 Volatility and correlation effects are small

In Section 2, we assume that the covariance matrix is fixed. We relax this assumption here. When the price of asset i changes, it may also affect volatilities and correlations. Considering these additional adjustments results in the inclusion of two additional terms in our demand elasticity decomposition, as shown in Proposition 2. The proof is provided in Appendix A.1.

Proposition 2. *The elasticity of MV investor demand can be decomposed into three parts:*

$$\eta_{i,t} = \underbrace{1 + \frac{1}{\mu_{i,unspanned,t}}\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right)}_{main\ component} + \underbrace{\left(\frac{1}{\sigma^2_{i,unspanned,t}}\right)\frac{\partial \sigma^2_{i,t}}{\partial \log(P_{i,t})}}_{volatility\ component} + \underbrace{\frac{1}{\mu_{i,unspanned,t}}\mu'_{-i,t}\frac{\partial \beta_{-i,t}}{\partial \log(P_{i,t})} - \left(\frac{1}{\sigma^2_{i,unspanned,t}}\right)\frac{\partial \sigma^2_{-i,t}}{\partial \log(P_{i,t})}}_{correlation\ component} \tag{13}$$

where $\frac{\partial \beta_{-i,t}}{\partial \log(P_{i,t})}$ *is an* $N-1$ *dimensional vector and the scalar* $\sigma^2_{-i,t} = \mathbb{V}ar(\beta'_{-i,t} r_{-i,t+1})$ *is the conditional variance of the replicating portfolio of asset i.*

In Equation (13), the first term is the same as in Proposition 1. The second term accounts for the effect of prices on the return volatility of asset i. The third term considers the effect on $\beta_{-i,t}$, the loading of asset i on other assets. The variance of the replicating portfolio, $\sigma^2_{-i,t}$, is included because it is influenced by $\beta_{-i,t}$, as $\sigma^2_{-i,t} = \mathbb{V}\mathrm{ar}_t(\beta'_{-i,t} r_{-i,t+1}) = \beta'_{-i,t}\Sigma_{-i,-i,t}\beta_{-i,t}$.

How large are these two additional terms? As discussed below, we estimate that incorporating these two effects changes the demand elasticity prediction by -2.2 and $+1.4$, respectively. Therefore, these effects are quantitatively small in explaining the substantial gap between theoretical and empirical estimates of demand elasticities, which is the central research question of this paper.

The volatility component. Accounting for volatility further decreases demand elasticity. Since Black (1976), it has been well-known that asset return volatility changes negatively with prices, a phenomenon often referred to as the "leverage effect." In Appendix B.3, consistent with the literature, we find that $\frac{\partial \sigma_{i,t}^2}{\partial \log(P_{i,t})}$ is approximately -0.06 for monthly return variance. Meanwhile, the average monthly $\sigma_{i,\text{unspanned},t}^2$ is approximately 0.027. Thus, this "volatility component" of demand elasticity is approximately $\frac{1}{0.027} \times (-0.06) \approx -2.2$. This, considering the volatility effect slightly lowers the demand elasticity prediction.

The correlation component. The correlation effect is analyzed by Davis (2024), who specifies a model where the stock covariance matrix contains both systematic and idiosyncratic components, with the systematic components parameterized by stock characteristics. When a stock's price changes, it affect the price-related characteristics such as the book-to-market ratio, thereby influencing the systematic component of the covariance matrix.[12] Davis (2024) empirically estimates the model's coefficients and finds that accounting for the correlation effect increases MV investor demand elasticity by approximately 1.4. Further details are provided in Appendix B.3.

[12]For example, if a stock's price declines, its book-to-market ratio increases, leading the stock to comove more with value stocks.

4.2 Consumption hedging effects are small

While we focus on MV demand, many researchers use demand derived from CRRA or Epstein-Zin (EZ) utility functions to incorporate wealth effects and intertemporal hedging considerations. We find these considerations have only minor effects on demand elasticities. This is intuitive, as we are examining the demand elasticity of an *individual* asset. As long as the investor holds a reasonably large number of assets, the relationship between the price of a single asset and future consumption should not be too large.

We follow Campbell, Chan, and Viceira (2003) to study EZ multivariate demand for N risky assets. Let y_t denote the N dimensional vector of log asset returns minus the log risk-free return. After log-linearization, the authors demonstrate that portfolio weights can be approximated by

$$w_t = \frac{1}{\tilde{\gamma}}\tilde{\Sigma}_t^{-1}\tilde{\mu}_t - \frac{\theta}{\psi\tilde{\gamma}}\tilde{\Sigma}_t^{-1}\sigma_{c-w,t}, \tag{14}$$

where $\tilde{\mu}_t = \mathbb{E}_t[y_{t+1}] + \frac{1}{2}\tilde{\sigma}_t^2$, $\tilde{\Sigma}_t$ is the $N \times N$ conditional covariance matrix of y_{t+1}, $\tilde{\sigma}_t^2$ is the N dimensional vector containing the diagonal elements of $\tilde{\Sigma}_t$, and $\sigma_{c-w,t}$ is the N dimensional vector of the conditional covariance between the log consumption-to-wealth ratio and y_{t+1}.[13] The preference parameters are as follows: $\tilde{\gamma} > 0$ is the relative risk aversion coefficient, $\psi > 0$ is the elasticity of intertemporal substitution, and $\theta \equiv (1-\tilde{\gamma})/(1-\psi^{-1})$. Note that CRRA is a special case of EZ with $\frac{\theta}{\psi\tilde{\gamma}} = 1$.

In Equation (14), the first term is identical to MV demand, aside from the quantitatively minor difference between using log returns versus simple returns (i.e., $\tilde{\Sigma}_t$ vs. Σ_t). The second term

[13] See equation (20) in Campbell et al. (2003). Their expression is more general and includes additional terms, as they consider y_t as the log return relative to a benchmark with possible additional covariance terms. In our case, the benchmark is the risk-free rate which simplifies the expression by omitting those extra terms.

introduces a new component that captures consumption-hedging considerations, which leads to an additional term in demand elasticity in Proposition 3 below. The proof is given in Appendix A.1.

Proposition 3. *The elasticity of demand for an EZ investor is given by:*

$$\eta_{i,t} = 1 + \underbrace{\frac{1}{\mu_{i,unspanned,t}}\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right)}_{\textit{main component}} + \underbrace{\frac{1}{\mu_{i,unspanned,t}} \cdot \frac{\theta}{\psi} \cdot \frac{\partial \sigma_{i,c-w,t}}{\partial \log(P_{i,t})}}_{\textit{consumption-hedging component}} \tag{15}$$

where $\sigma_{i,c-w,t}$ is the covariance between asset i return and the log consumption-to-wealth ratio.

If assets have stable covariance with the consumption-to-wealth ratio, then the consumption hedging component would be zero. It can only play a role if the covariance varies with prices. To assess the significance of the component, we provide a calibration, with details provided in Appendix B.4. We find the effect to be quantitatively small. Specifically, we aim to calculate the sensitivity of $\mathbb{E}_t[r^{unex}_{t+1} cay_{t+1}]$ to price changes, where $r^{unex}_{t+1} \equiv r_{t+1} - \mathbb{E}_t[r_{t+1}]$ and cay_{t+1} represents the consumption-to-wealth ratio log deviations from Lettau and Ludvigson (2001). Since the *cay* data is available quarterly, we use quarterly returns. To approximate r^{unex}_{t+1}, we take the residuals from a regression of quarterly returns on the van Binsbergen et al. (2023) price wedge used earlier, including stock fixed effects. We then regress $r^{unex}_{t+1} \cdot cay_{t+1}$ on the van Binsbergen et al. (2023) price instrument, again including stock fixed effects, which produces a slope coefficient of -0.0005. Standard errors are double-clustered at the stock and quarter level, yielding a standard error of 0.0004 (Table B.3 in the appendix). This indicates the hedging component of demand elasticity is not statistically different from zero. When translated to a monthly level, our estimate indicates that $\frac{\partial \sigma_{i,c-w,t}}{\partial \log(P_{i,t})}$ is around -0.00017 ($\approx -0.0005/3$), which is two orders of magnitude smaller than our pass-through estimates of 0.014. This suggests that the EZ hedging component affects demand elasticity by about 1% of its total value, according to the point estimate.

4.3 Transaction costs

In Section 3, we show that empirically estimated models of stock returns reduce MV demand elasticity to around 7. While this is significantly lower than the theoretical calibrations in the thousands, it remains higher than empirical estimates. For instance, Koijen, Richmond, and Yogo (2024) estimate a demand elasticity of about 0.5 for the hedge fund sector. To address this "remaining gap," we explore one potential factor: transaction costs, which can significantly lower elasticity. However, we acknowledge that many alternative mechanisms may be simultaneously at play, and we do not take a stance on which is the most relevant.[14]

In this section, we demonstrate that applying cost optimization to MV demand, following the approach in Gârleanu and Pedersen (2013), can reduce demand elasticity from 7 to levels consistent with empirical estimates. We explain the intuition behind the optimizer, present a simple equation for cost-optimized elasticity, and provide a rough calibration suggesting that cost optimization may bring elasticity in line with the empirical estimates. Further details are provided in Appendix B.5.

In the cost optimization model of Gârleanu and Pedersen (2013), the optimized portfolio weights, $w^*_{i,t}$, are a convex combination of the passively adjusted previous weights and an "aim" portfolio, $\text{aim}_{i,t}$, weighted by $s_{\text{aim}} \in [0,1]$. A higher s_{aim} indicates that the portfolio is weighted more heavily toward the aim portfolio and puts lower weights on the passive portfolio. As shown in Equation (B.5) in the appendix, the aim portfolio averages current and expected future non-cost-optimized weights, $\mathbb{E}_t[w_{t+\tau}]$, using geometric weights based on $\rho_{\text{aim}} \in [0,1]$. The future weights, $\mathbb{E}_t[w_{t+\tau}]$, have their own price elasticity that follows an AR(1) process with the parameter ρ_η (see Equation (B.8) in the appendix). Intuitively, a small s_{aim} leads to a predominantly passive,

[14] We provide a possible mechanism here for completeness, though readers less focused on this detail can proceed without significant impact on understanding our main findings.

inelastic portfolio. Even with a large s_{aim}, a high ρ_{aim} results in inelasticity as the portfolio targets a long-term position that is less sensitive to current prices.

Appendix B.5 shows that the cost-optimized elasticity $\eta^*_{i,t}$ is approximately:

$$\eta^*_{i,t} \approx s_{\text{aim}} \left(\frac{\text{aim}_{i,t}}{w^*_{i,t}} \right) \left(\frac{1 - \rho_{\text{aim}}}{1 - \rho_\eta \rho_{\text{aim}}} \right) \eta_{i,t}. \tag{16}$$

where ρ_η is a parameter slightly lower than one and explained further in Appendix B.5. In our calibration in the Appendix, cost optimization reduces elasticity to within the empirical range of 0.3 to 1.6, as reported by Gabaix and Koijen (2022). Table 4 summarizes the parameter values and the resulting cost-optimized elasticity. We set $\rho_\eta = 0.99$ (with lower values further reducing elasticity) and consider s_{aim} between 0.03 and 0.1, and ρ_{aim} between 0.9 and 0.97. For the non-cost-optimized elasticity, $\eta_{i,t}$, we use a value of 7, as previously discussed, and also explore higher values of 20, 50, and 1,200 to assess the impact of increased elasticity. We consider $\text{aim}_{i,t}/w^*_{i,t} \in [0.5, 2]$, meaning the aim portfolio targets anywhere from doubling to halving the current position. Table 4 shows that with $\eta_{i,t} = 7$, cost optimization brings elasticity within the empirical range, though it remains above this range for the high unspanned return elasticity of 1,200.

In summary, low pass-throughs, high unspanned returns, and trading costs are sufficient to deliver inelastic demand consistent with the empirical estimates. It is important to note that many other mechanisms can also reduce demand elasticities. In practice, institutional investors face leverage constraints, short-selling constraints, and are often subject to benchmarking or indexing incentives (e.g., Basak and Pavlova, 2013; Koijen and Yogo, 2019; Haddad et al., 2022). All of these factors may contribute to lower demand elasticities. Therefore, while transaction costs are one possible explanation, we do not take a stance on which mechanisms are the most relevant.

s_{aim}	ρ_{aim}	$\eta_{i,t}$	Cost Optimized Elasticity ($\eta^*_{i,t}$)	s_{aim}	ρ_{aim}	$\eta_{i,t}$	Cost Optimized Elasticity ($\eta^*_{i,t}$)
0.03	0.97	7	[0.1, 0.3]	0.10	0.97	7	[0.3, 1.1]
		20	[0.2, 0.9]			20	[0.8, 3.0]
		50	[0.6, 2.3]			50	[1.9, 7.6]
		1200	[13.6, 54.4]			1200	[45.3, 181.4]
0.03	0.90	7	[0.1, 0.4]	0.10	0.90	7	[0.3, 1.3]
		20	[0.3, 1.1]			20	[0.9, 3.7]
		50	[0.7, 2.8]			50	[2.3, 9.2]
		1200	[16.5, 66.1]			1200	[55.0, 220.2]

Table 4. Cost optimized demand elasticity
This table shows the cost-optimized demand elasticity ranges using Equation (16) with several values for parameters s_{aim} and ρ_{aim}. We provide a range for the cost-optimized elasticity, $\eta^*_{i,t}$, because we use $\text{aim}_{i,t}/w^*_{i,t} \in [0.5, 2]$. In our calibration, we use $\rho_\eta = 0.99$.

5 Understanding High Unspanned Returns

Our analysis thus far suggests that unspanned returns are elevated, which implies that demand elasticities should be low. In finance, factor models are traditionally employed to explain the cross-section of expected returns. Within this framework, one might question how high unspanned returns fit within the factor model context. At first glance, our finding seems at odds: if factor models can accurately price assets, shouldn't unspanned returns be zero? Furthermore, it is also worth checking whether the high unspanned returns observed in our empirical results are unique to the model examined in Section 3.

In this section, we address these two questions. First, we theoretically show that, even if expected returns are fully captured by a factor model, unspanned returns remain large if the model exhibits "weak factors," as defined by Lettau and Pelger (2020). Second, we empirically implement a broad set of asset pricing models from the literature. We show that, once these models are allowed to incorporate weak factors, they all predict high unspanned returns.

5.1 Theoretical results

This section introduces a proposition that helps interpret unspanned returns within the context of factor models.

Unspanned returns as "weak factors". We drop the t subscripts for simplicity. Consider the pricing of N securities, where r represents the vector of realized excess returns, and $\mu = \mathbb{E}(r)$ and $\Sigma = \mathbb{V}\mathrm{ar}(r)$ denote the expected return and the covariance matrix, respectively. Consider a factor model with F factors, where the factor portfolio weights are given by an $N \times F$ matrix W. Thus, the realized and expected factor excess returns are defined as $f = W'r$ and $\mu_f = W'\mu$, respectively. The vector of expected excess returns for assets can then be expressed as:

$$\mu = \alpha + B \cdot \mu_f,$$

where α is a vector of pricing errors, and $B = \Sigma W(W'\Sigma W)^{-1}$ is the $N \times F$ matrix of factor loadings. Because we allow for alphas, or pricing errors, this is without loss of generality. Additionally, we define $\omega^f = \mathbb{V}\mathrm{ar}^{-1}(f) \cdot \mathbb{E}(f)$, which represents the vector of (unscaled) factor portfolio weights for a mean-variance efficient investor. Proposition 4 connects the stock-level unspanned returns to the factor model.

Proposition 4. *The unspanned return of asset i can be decomposed into two terms:*

$$\mu_{i,unspanned} = \underbrace{\alpha_i - \beta'_{-i}\alpha_{-i}}_{\alpha_{i,unspanned}} + \underbrace{\Sigma'_i(W - W^*_{-i})\omega^f}_{\textit{weak factor term for asset } i}, \tag{17}$$

where Σ_i is the i^{th} column of the covariance matrix Σ and

$$W_{-i}^{*} = \underbrace{\left(I_{-i}'\Sigma_{-i,-i}^{-1}I_{-i}\Sigma\right)}_{\text{projection-like matrix}} W, \tag{18}$$

where I_{-i} is the identity matrix with the i^{th} row removed (resulting in an $(N-1)\times N$ matrix), and $\Sigma_{-i,-i}$ is the $(N-1)\times(N-1)$ covariance matrix excluding asset i.

Proof. See Appendix A.1. □

Proposition 4 states that, in the context of a factor model, the unspanned return of an asset is determined by its unspanned alpha in the first term and its association with "weak factors" in the second term.

The first term in Equation (17) represents the alpha of asset i minus the alpha of its replicating portfolio, formed using all other assets. This part is intuitive: if the factor model does not explain expected returns well, unspanned returns can be high.

The second term in Equation (17) represents the more important realization: even if factor model alphas are zero, we can still have non-zero unspanned returns. Intuitively, this term captures the *incremental* contribution of asset i to the factor model, as the term $W - W_{-i}^{*}$ represents the difference in factor weights with and without asset i. Specifically, the matrix W_{-i}^{*} represents the factor portfolio weights that drop any weight on asset i by placing a zero in its place and update weights on other assets based on the importance of asset i. This importance is calculated using the projection-like matrix, $\left(I_{-i}'\Sigma_{-i,-i}^{-1}I_{-i}\Sigma\right)$. The $(N-1)\times(N-1)$ matrix $\Sigma_{-i,-i}^{-1}$ is the inverse of the covariance matrix without asset i. The matrix $I_{-i}'\Sigma_{-i,-i}^{-1}I_{-i}$ expands this back up to an $N\times N$ matrix, by placing zeros in all positions related to asset i. Thus, this matrix can be viewed as the inverse covariance matrix that excludes asset i. If asset i is not important in the covariance matrix, then $I'_{-i}\Sigma^{-1}_{-i,-i}I_{-i}\Sigma \approx I$, which implies $W^*_{-i} \approx W$, and the weak factor term $\Sigma'_i(W - W^*_{-i})\omega^f$ will be very close to zero. In the language of Lettau and Pelger (2020), the second term in Equation (17) captures the "weak factor" term associated with asset i, and its size is determined by the importance of asset i for the factor model.

Can an individual asset play a non-negligible role in a factor model? This depends on the nature of expected returns. Classical models such as the CAPM assume that only factors that explain a large amount of time-series variation matter for expected returns, while other factor portfolios have zero expected excess returns. If this were the case, removing a single asset from the factor model would have a negligible effect, consistent with the calibration in Section 2.3.[15] However, in recent decades, researchers have identified an increasing number of "weak factors" that can explain non-zero variation in the cross-section of expected returns, even if they do not account for a high fraction of time-series return variation (e.g., Lettau and Pelger, 2020). In such cases, the second term in Equation (17) can become non-negligible.

5.2 Empirical illustrations

In this section, we examine the demand elasticities implied by a wide range of factor models. This analysis illustrates the central intuition from Section 5.1, which suggests that factor models with more weak factors exhibit higher unspanned returns and, consequently, lower demand elasticities. To isolate the effect of unspanned returns, we apply the same price pass-through of 0.014 from Section 2 when calculating demand elasticities across all models. To compute unspanned returns

[15] We can use Proposition 4 to arrive at an identical model-implied elasticity as in Section 2.3. Consider a CAPM model, where W represents the market portfolio weights. As before, assume monthly pass-through is equal to one, and all $N = 1{,}000$ stocks have an annual return of 6%, volatility of 30%, and a correlation of 0.3. The market portfolio weights are simply $W = \frac{1}{N} \cdot \mathbf{1}$, where $\mathbf{1}$ is a vector of ones. We can compute the projection-like matrix in (18), and find that $W^*_{-i} = (0.001001, \ldots, 0.001001)'$. Then, from (17), the unspanned return is $\mu_{\text{unspanned}} \approx 0.0012\%$, leading to an elasticity of about 7,000, identical to what we find in Section 2.3.

of any factor model, we need three components: unspanned variance, factor portfolio weights, and the weight of the MV efficient portfolio in each factor (ω_f). Appendix B.6 provides further details.

We implement two types of factor models: classical models (e.g., CAPM, FF3, etc.) and those based on a larger number of characteristics that utilize principal component analysis (PCA). The details of each type are discussed below.

Classical factor models. We begin by considering classical factor models with an increasing number of factors. These include the CAPM, Fama and French (1993) (FF3), Hou, Xue, and Zhang (2015) (HXZ), Fama and French (2015) (FF5), and Fama and French (2015) with the momentum factor (FF6). For model implementation, we follow the methodology outlined in the original papers to construct the factor portfolio weights. For example, Fama and French (1993) describe how they construct their $N \times 3$ matrix of portfolio weights W_t, where the column of market portfolio weights is simply the market equity divided by the sum of market equity across all stocks. The high-minus-low factor portfolio weights are calculated as the average of the "Small Value" and "Big Value" weights minus the average of the "Small Growth" and "Big Growth" portfolio weights, and so on.

We then construct the stochastic discount factor (SDF) portfolio weight in factors, ω^f, using its empirical counterpart as in Brandt, Santa-Clara, and Valkanov (2009): $\widehat{\mathbb{V}\mathrm{ar}}^{-1}(f)\hat{\mathbb{E}}(f)$, where $\widehat{\mathbb{V}\mathrm{ar}}(\cdot)$ and $\hat{\mathbb{E}}(\cdot)$ are the empirical estimates of the covariance matrix and expected value vector, respectively.[16]

To calculate the residual variance, we model the covariance matrix of stock returns as $\Sigma_t = \beta_t \Omega \beta_t' + E_t$, where Ω is the covariance matrix of factor returns estimated from the full sample, β_t are time-varying betas, and E_t is a diagonal matrix containing the unspanned variances. We estimate betas using five-year rolling window time-series regressions. We then estimate the diagonal terms

[16] Kozak, Nagel, and Santosh (2020) discuss an extension of Brandt et al. (2009) where shrinking is applied, though it has minimal impact on our results. Davis (2024) outline alternative ways of calculating ω^f.

of E_t using five-year rolling windows, after subtracting the factor-explained components from stock returns.

PCA factor models. Recently, researchers have proposed several approaches for constructing models with PCA-style factors based on a broad set of stock characteristics. These models primarily differ in how they form characteristics-based factors. To implement these models, we use data for $K = 62$ stock characteristics from Freyberger et al. (2020) and transform them into uniform distributions between -0.5 and 0.5 using cross-sectional ranks, following a normalization procedure employed by Kelly et al. (2019) and Kozak et al. (2020), among others. In the model descriptions below, Z_t denotes the $N \times K$ matrix where these stock characteristics are stacked.

In these models, the number of factors can be specified by the researcher, and we implement all possible configurations. The initial factors selected typically explain a larger fraction of time-series return variation, similar to those in classical models, while the later factors tend to be progressively weaker. We consider three types of PCA factor models:

1. **RPCA.** Chen et al. (2023) develop a method where characteristic-based portfolio returns are obtained by regressing returns, r_{t+1}, on lagged characteristics, Z_t: $y_{t+1} = (Z_t' Z_t)^{-1} Z_t' r_{t+1}$. Regressed PCA (RPCA) factors are then calculated through a PCA decomposition on the covariance matrix of y_{t+1}. To calculate the residual variance for stocks, we model the stock return covariance matrix as $\Sigma_t = \beta_t \Omega \beta_t' + \sigma_\epsilon^2 I$, where Ω represents the covariance matrix of factor returns, β_t is the vector of betas estimated from Chen et al. (2023), and σ_ϵ^2 is the average stock-level residual variance.

2. **RP-PCA.** We also implement the risk-premium PCA (RP-PCA) model in Lettau and Pelger (2020). Unlike traditional PCA models, which construct factors based solely on the covariance matrix of time-series variation (e.g., Connor and Korajczyk, 1986), the RP-

PCA model also accounts for the cross-section of returns. As explained in Lettau and Pelger (2020), this model is designed to capture "weak factors" that explain cross-sectional variation, even if they have limited contribution to time-series variation. Following their baseline specification, we set γ, the parameter in Lettau and Pelger (2020) that governs the weight assigned to cross-sectional variation, to 10. To ensure consistency with the other PCA models, we apply their procedure to portfolios formed through linear projection onto characteristics (which generates the y_{t+1} portfolio returns above), rather than portfolios formed by sorting based on characteristics, as in the original paper. We use the same method for calculating the covariance matrix and residual variance as with the RPCA model.

3. **IPCA.** Finally, we implement the instrumented PCA (IPCA) model from Kelly et al. (2019). Their methodology produces estimates for factor portfolio weights and betas as by-products of the optimization process. For further details, we refer the reader to their original paper. We use the same method to calculate the covariance matrix and the residual variance as for the RPCA model.

To focus on the role of weak factors, we implement versions of these three PCA-style factor models without alpha terms. In other words, we assume that the factors fully explain the cross-section of expected returns. The results that include alphas are presented in Appendix C.4. Consistent with Proposition 4, when alphas are included, unspanned returns increase, and demand elasticities decline further.

For each of these PCA factor models, we compute demand elasticities as a function of the number of factors included. Following the procedure outlined in Section 3, we calculate the average unspanned return conditional on having a positive MV portfolio weight. Appendix B.6 provides additional details on how unspanned returns are computed. We then compute the implied demand elasticity, assuming a pass-through of 0.014, as estimated in Section 3.2.

Table 5 presents the results. Classical factor models exhibit extremely elastic demand, consistent with the notion that these models imply assets are highly substitutable, as returns are driven by only a few systematic factors. In contrast, the data-driven PCA factor models display lower demand elasticities, consistent with the idea that stocks are imperfect substitutes for one another.

In Figure 3, we plot the demand elasticity of these asset pricing models as the number of factors increases. Among the classical factor models, the CAPM exhibits the highest demand elasticity, reaching into the thousands, consistent with the discussion in Section 2.3. Notably, as additional factors are incorporated, demand elasticity steadily declines, reflecting larger unspanned returns. The FF3 model, which adds size and value factors to the CAPM, already shows demand elasticity below one thousand, while the FF5 and FF6 models have elasticities around one hundred.

Compared to the classical factor models, the PCA factor models exhibit lower demand elasticity, particularly as the number of factors increases. Among the three PCA models, RPCA shows the highest demand elasticity, followed by RP-PCA, with IPCA exhibiting the lowest, around 20 with just four factors. Importantly, as additional (weaker) factors are included, the demand elasticity of all three PCA models steadily declines, converging to approximately 7 when the number of factors approaches the number of characteristics.[17] Interestingly, this is similar to the demand elasticity we estimated in Section 3.4.

To summarize, this section shows that high unspanned returns are consistent with the cross-section of expected stock returns explained by factor models that include weak factors. Additionally, we empirically implement several recently proposed factor models. When these models include weak factors, they consistently exhibit high unspanned returns and predict low demand elasticities, in line with our earlier findings in Section 3. This suggests that our previous estimation is not unique to the specific model used but is a common feature of factor models with weak factors.

[17]It is important to note that we are not forcing these models to include weak factors. If the true model only contains K factors, even if we allow for $F > K$ factors, the estimation will show that the factors $K + 1, \ldots, F$ have little variation.

Factor Model	Number of Factors	Unspanned Return	Weight Responsiveness	Demand Elasticity
Classical Factor Models				
CAPM	1	0.0006%	159678	2236.5
FF3	3	0.0020%	51145	717.0
HXZ	4	0.0083%	12018	169.3
FF5	5	0.0124%	8063	113.9
FF6	6	0.0109%	9164	129.3
RPCA				
PCA Factor Model	5	0.0207%	4839	68.7
PCA Factor Model	10	0.0313%	3193	45.7
PCA Factor Model	15	0.0753%	1328	19.6
PCA Factor Model	20	0.0795%	1258	18.6
PCA Factor Model	50	0.1956%	511	8.2
PCA Factor Model	60	0.2449%	408	6.7
RP-PCA				
PCA Factor Model	5	0.0368%	2716	39.0
PCA Factor Model	10	0.0549%	1820	26.5
PCA Factor Model	15	0.0830%	1204	17.9
PCA Factor Model	20	0.1026%	975	14.6
PCA Factor Model	50	0.1920%	521	8.3
PCA Factor Model	60	0.2373%	421	6.9
IPCA				
PCA Factor Model	5	0.0886%	1129	16.8
PCA Factor Model	10	0.1112%	900	13.6
PCA Factor Model	15	0.1431%	699	10.8
PCA Factor Model	20	0.1560%	641	10.0
PCA Factor Model	50	0.2085%	480	7.7
PCA Factor Model	60	0.2414%	414	6.8

Table 5. Demand Elasticity Across Factor Models
This table presents the demand elasticity, calculated using Equation (9), for various factor models. The first set consists of classical factor models. The next three sets are PCA-style models, where RPCA refers to the regression PCA model in Chen et al. (2023), RP-PCA refers to the risk-premium PCA model in Lettau and Pelger (2020), and IPCA is the instrumented PCA model in Kelly et al. (2019). For the PCA factor models, we vary the number of factors included. The second column reports the average monthly unspanned returns for stocks with positive MV portfolio weights. The third column reports the weight responsiveness, calculated as the reciprocal of unspanned returns (Equation 8). The elasticity is computed with a pass-through of 0.014 using Equation (9). Appendix B.6 offers further details on how we compute the unspanned returns of these models.

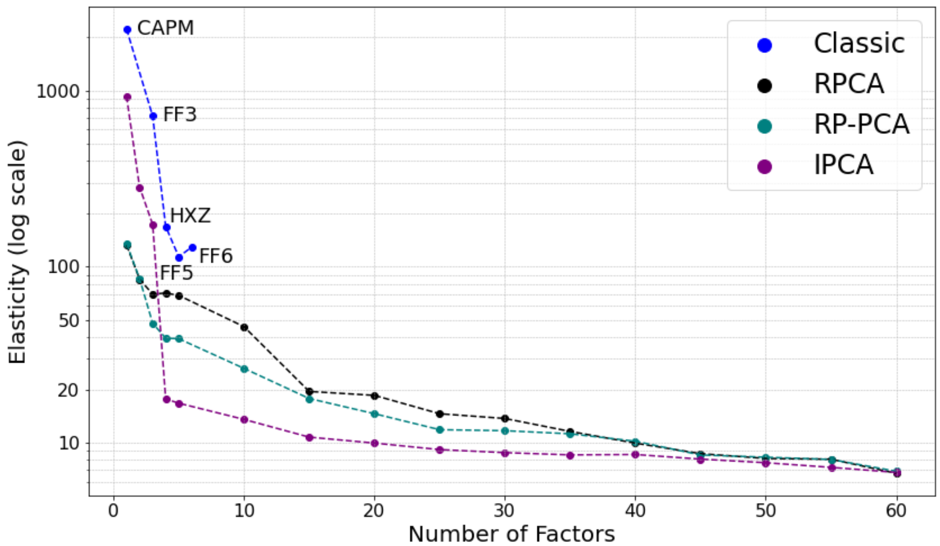

Figure 3. Demand Elasticity Across Factor Models
This figure examines a variety of factor models with different numbers of factors. The x-axis displays the number of factors, while the model-implied demand elasticity is plotted on a log scale on the y-axis, using a price pass-through of 0.014. In the legend, "Classic" refers to traditional factor models. RPCA refers to the model in Chen et al. (2023), RP-PCA refers to the model in Lettau and Pelger (2020), and IPCA refers to the model in Kelly et al. (2019).

6 Conclusion

Classical asset pricing theories predict extremely high stock-level demand elasticities, often in the thousands, implying that trading flows would have minimal impact on asset prices. However, empirical studies find demand elasticities around one, which is three orders of magnitude lower. In this paper, we show that this gap is largely explained by incorporating empirically estimated properties of stock returns. For investors forming mean-variance efficient portfolios, using empirically estimated—rather than assumed—moments of stock returns reduces the predicted demand elasticity from 7,000 to around 7.

We begin the analysis by analytically showing that mean-variance investor demand elasticity decomposes into two components: price pass-through and the reciprocal of unspanned returns. The former measures how much cash flow-unrelated price movements predict future returns, while the latter captures the extent to which assets are substitutable. The less substitutable an asset is, the less its returns are spanned by others, leading to lower demand elasticity.

Guided by this analytical result, we then empirically estimate these two components for U.S. stocks. We estimate the price pass-through to be around 0.014, meaning that a 1% cash flow-unrelated price drop leads to an increase in next-month return of 1.4 basis points. Using this empirically estimated pass-through, rather than classical theoretical assumptions, reduces the predicted demand elasticity from 7,000 to 1,200. More importantly, when estimating unspanned returns, we find that stocks are far from the perfect substitutes typically assumed in theory. Incorporating empirically estimated unspanned returns further reduces the demand elasticity to around 7. We show that stocks exhibiting high unspanned returns can be understood as exhibiting "weak factors", and our finding that stocks are imperfect substitutes is consistent with the literature showing that the cross-section of stock returns is poorly spanned by systematic risk factors.

References

Baba Yara, Fahiz, Brian H. Boyer, and Carter Davis, 2021, The factor model failure puzzle, Working Paper, Indiana.

Bartram, Söhnke M, and Mark Grinblatt, 2018, Agnostic fundamental analysis works, *Journal of Financial Economics* 128, 125–147.

Basak, Suleyman, and Anna Pavlova, 2013, Asset prices and institutional investors, *American Economic Review* 103, 1728–1758.

Black, Fischer, 1976, Studies of stock price volatility changes, in *Proceedings of the Business and Economic Statistics Section*, 177–181.

Bouchaud, Jean-Philippe, Julius Bonart, Jonathan Donier, and Martin Gould, 2018, *Trades, Quotes and Prices: Financial Markets Under the Microscope* (Cambridge University Press).

Brandt, Michael W, 2010, Portfolio choice problems, in *Handbook of Financial Econometrics: Tools and Techniques*, 269–336 (Elsevier).

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447.

Campbell, John Y., Yeung Lewis Chan, and Luis M. Viceira, 2003, A multivariate model of strategic asset allocation, *Journal of Financial Economics* 67, 41–80.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *Review of Financial Studies* 28, 212–246.

Chen, Qihui, Nikolai Roussanov, and Xiaoliang Wang, 2023, Semiparametric conditional factor models: Estimation and inference, Working Paper 31817, National Bureau of Economic Research.

Connor, Gregory, and Robert A Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, *Journal of financial economics* 15, 373–394.

Davis, Carter, 2024, Elasticity of quantitative investment, *Review of Financial Studies* forthcoming.

De Bondt, Werner F.M., and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

Dello-Preite, Massimo, Raman Uppal, Paolo Zaffaroni, and Irina Zviadadze, 2024, Cross-sectional asset pricing with unsystematic risk, Working paper.

Duffie, Darrell, 2010, Presidential address: Asset price dynamics with slow-moving capital, *The Journal of finance* 65, 1237–1267.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1 – 22.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2018, Trading costs, Working Paper, AQR and Yale.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2020, Dissecting characteristics nonparametrically, *Review of Financial Studies* 33, 2326–2377.

Gabaix, Xavier, and Ralph S.J. Koijen, 2022, In search of the origins of financial fluctuations: The inelastic markets hypothesis, Working Paper 28967, National Bureau of Economic Research.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2013, Dynamic trading with predictable returns and transaction costs, *Journal of Finance* 68, 2309–2340.

Gromb, Denis, and Dimitri Vayanos, 2010, Limits of arbitrage, *Annu. Rev. Financ. Econ.* 2, 251–275.

Haddad, Valentin, Paul Huebner, and Erik Loualiche, 2022, How competitive is the stock market? Theory, evidence from portfolios, and implications for the rise of passive investing, Working Paper, UCLA and Minnesota.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *The Review of Financial Studies* 28, 650–705.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

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.

Kim, Soohun, Robert A Korajczyk, and Andreas Neuhierl, 2021, Arbitrage portfolios, *Review of Financial Studies* 34, 2813–2856.

Koijen, Ralph SJ, Robert J Richmond, and Motohiro Yogo, 2024, Which investors matter for equity valuations and expected returns?, *Review of Economic Studies* 91, 2387–2424.

Koijen, Ralph S.J., and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross section, *Journal of Financial Economics* 135, 271 – 292.

Ledoit, Olivier, and Michael Wolf, 2004, Honey, I shrunk the sample covariance matrix, *Journal of Portfolio Management* 30, 110–119.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

Lettau, Martin, and Markus Pelger, 2020, Factors that fit the time series and cross-section of stock returns, *Review of Financial Studies* 33, 2274–2325.

Lewellen, Jonathan, 2015, The cross section of expected stock returns, *Critical Finance Review* 4, 1–44.

Lopez-Lira, Alejandro, and Nikolai L. Roussanov, 2023, Do common factors really explain the cross-section of stock returns?, Working Paper, Florida and Wharton.

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

Modigliani, Franco, and Merton H Miller, 1958, The cost of capital, corporation finance and the theory of investment, *The American economic review* 48, 261–297.

Pavlova, Anna, and Taisiya Sikorskaya, 2023, Benchmarking intensity, *Review of Financial Studies* 36, 859–903.

Petajisto, Antti, 2009, Why do demand curves for stocks slope down?, *Journal of Financial and Quantitative Analysis* 44, 1013–1044.

Schmickler, Simon, 2020, Identifying the price impact of fire sales using high-frequency surprise mutual fund flows, Working Paper, Princeton.

Scholes, Myron S, 1972, The market for securities: Substitution versus price pressure and the effects of information on share prices, *The Journal of Business* 45, 179–211.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *Journal of Finance* 41, 579–590.

Stevens, Guy VG, 1998, On the inverse of the covariance matrix in portfolio analysis, *Journal of Finance* 53, 1821–1827.

van Binsbergen, Jules H., Martijn Boons, Christian C. Opp, and Andrea Tamoni, 2023, Dynamic asset (mis) pricing: Build-up versus resolution anomalies, *Journal of Financial Economics* 147, 406–431.

van der Beck, Philippe, 2022, On the estimation of demand-based asset pricing models, Working Paper, HBS.

Appendix

A Supporting Derivations

A.1 Proofs of Propositions

Proof of Proposition 1. In this proof, we drop the t subscripts and tildes for notational simplicity. Without loss of generality, just consider the last asset, asset N. Subdivide the matrix into blocks

$$\Sigma = \begin{bmatrix} \Sigma_{-N,-N} & \Sigma_{-N} \\ \Sigma'_{-N} & \sigma_N^2 \end{bmatrix}.$$

Using the block diagonal matrix formula, note that:

$$\Sigma^{-1} = \begin{bmatrix} \Sigma^{-1}_{-N,-N} + \tau_N \Sigma^{-1}_{-N,-N} \Sigma_{-N} \Sigma'_{-N} \Sigma^{-1}_{-N,-N} & -\tau_N \Sigma^{-1}_{-N,-N} \Sigma_{-N} \\ -\tau_N \Sigma'_{-N} \Sigma^{-1}_{-N,-N} & \tau_N \end{bmatrix}$$

where $\tau_N = (\sigma_N^2 - \Sigma'_{-N} \Sigma^{-1}_{-N,-N} \Sigma_{-N})^{-1}$. The MV optimal portfolio weights are (up to a multiplicative constant) given by:

$$\begin{aligned} & w = \Sigma^{-1}\mu \\ \Rightarrow & \begin{cases} w_{-N} & = \left(\Sigma^{-1}_{-N,-N} + \tau_N \Sigma^{-1}_{-N,-N} \Sigma_{-N} \Sigma'_{-N} \Sigma^{-1}_{-N,-N} \right) \mu_{-N} - \tau_N \Sigma^{-1}_{-N,-N} \Sigma_{-N} \cdot \mu_N \\ \\ w_N & = -\tau_N \Sigma'_{-N} \Sigma^{-1}_{-N,-N} \mu_{-N} + \tau_N \mu_N \end{cases} \end{aligned} \tag{A.1}$$

To make these expressions more intuitive, note that $\beta_{-N} = \Sigma^{-1}_{-N,-N} \Sigma_{-N}$ is the $N-1$ vector of the beta of asset N to all other assets. Therefore, we can simplify (A.1) into:

$$\begin{aligned} w_{-N} &= (\Sigma^{-1}_{-N,-N} + \tau_N \beta_{-N} \beta'_{-N}) \mu_{-N} - \tau_N \beta_{-N} \mu_N \\ w_N &= -\tau_N \underbrace{\beta'_{-N} \mu_{-N}}_{= \mu_{N,\text{spanned}}} + \tau_N \mu_N \end{aligned} \tag{A.2}$$

These expressions make clear the "hedging relationship": if μ_N changes, the investor responds by increasing holdings w_N but also reduces w_{-N} in a way that is proportional to β_{-N}, the portfolio that hedges N using all other assets. This can be thought of as an arbitrage trade with asset N on the long side and the other assets on the short side (with portfolio weights β_{-N}). Also, $\tau_N^{-1} = \sigma_N^2 - \beta'_{-N}\Sigma_{-N,-N} \cdot \beta_{-N}$ is the residual variance of N after hedging out exposure to other assets.

We are now ready to derive Equation (8) in Proposition 1. Rewrite Equation (A.2) and take derivatives:

$$w_N = \tau_N \cdot \left(\mu_N - \underbrace{\beta'_{-N}\mu_{-N}}_{=\mu_{N,\text{spanned}}}\right) = \tau_N \cdot \mu_{N,\text{unspanned}}.$$

$$\Rightarrow \theta_N = \frac{\partial \log(w_N)}{\partial \mu_N} = \frac{1}{w_N} \cdot \frac{\partial w_N}{\partial \mu_N} = \frac{1}{\mu_{N,\text{unspanned}}}.$$

Also, note that:

$$\tau_N = \frac{1}{\sigma^2_{N,\text{unspanned}}} \Rightarrow w_N = \frac{\mu_{N,\text{unspanned}}}{\sigma^2_{N,\text{unspanned}}},$$

which also shows that Equation (10) holds. □

Proof of Proposition 2. We take the derivative of Equation (10) as a function of prices:

$$\begin{aligned}
\eta_{i,t} &= 1 - \frac{1}{w_{i,t}} \cdot \frac{\partial w_{i,t}}{\partial \log(P_{i,t})} \\
&= 1 - \frac{1}{w_{i,t}} \left(\frac{\partial w_{i,t}}{\partial \mu_{i,t}} \frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})} + \left(\frac{\partial w_{i,t}}{\partial \beta_{-i,t}}\right)' \frac{\partial \beta_{-i,t}}{\partial \log(P_{i,t})} + \frac{\partial w_{i,t}}{\partial \sigma^2_{i,\text{unspanned},t}} \frac{\partial \sigma^2_{i,\text{unspanned},t}}{\partial \log(P_{i,t})} \right) \\
&= 1 + \frac{1}{w_{i,t}} \left[\frac{1}{\gamma A_t \sigma^2_{i,\text{unspanned},t}} \left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right) + \left(\frac{\mu_{-i,t}}{\gamma A_t \sigma^2_{i,\text{unspanned},t}}\right)' \frac{\partial \beta_{-i,t}}{\partial \log(P_{i,t})} \right. \\
&\quad \left. + \left(\frac{\mu_{i,t} - \beta'_{-i,t}\mu_{-i,t}}{\gamma A_t \sigma^4_{i,t,\epsilon}}\right) \frac{\partial \sigma^2_{i,\text{unspanned},t}}{\partial \log(P_{i,t})} \right] \\
&= 1 + \frac{1}{\mu_{i,\text{unspanned},t}} \left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})} + \mu'_{-i,t} \frac{\partial \beta_{-i,t}}{\partial \log(P_{i,t})} \right) + \left(\frac{1}{\sigma^2_{i,\text{unspanned},t}}\right) \frac{\partial \sigma^2_{i,\text{unspanned},t}}{\partial \log(P_{i,t})}, \quad \text{(A.3)}
\end{aligned}$$

where the last step uses the result from Proposition 1 that $\mu_{i,\text{unspanned},t} = w_{i,t}\gamma A_t \sigma^2_{i,\text{unspanned},t}$. Here, $\partial w_{i,t}/\partial \beta_{-i,t}$ and $\partial \beta_{-i,t}/\partial \log(P_{i,t})$ are $(N-1)$ dimensional vectors, while the other terms are scalars.

Using Equation (6), we can write $\sigma^2_{i,\text{unspanned},t} = \sigma^2_{i,t} - \sigma^2_{-i,t}$, where $\sigma^2_{-i,t}$ is the scalar conditional variance of the replicating portfolio, $\sigma^2_{-i,t} = Var_t(\beta'_{-i,t} r_{-i,t+1})$. Plugging this into Equation (A.3) gives:

$$\eta_{i,t} = \underbrace{1 + \frac{1}{\mu_{i,\text{unspanned},t}}\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right)}_{\text{main component}}$$
$$+ \underbrace{\frac{1}{\mu_{i,\text{unspanned},t}}\mu'_{-i,t}\frac{\partial \beta_{-i,t}}{\partial \log(P_{i,t})} - \left(\frac{1}{\sigma^2_{i,\text{unspanned},t}}\right)\frac{\partial \sigma^2_{-i,t}}{\partial \log(P_{i,t})}}_{\text{correlation component}} + \underbrace{\left(\frac{1}{\sigma^2_{i,\text{unspanned},t}}\right)\frac{\partial \sigma^2_{i,t}}{\partial \log(P_{i,t})}}_{\text{Black leverage component}}.$$

□

Proof of Proposition 3. To focus on the demand for asset i, we modify Proposition 1 by using demand of Equation (14) in Equation (10):

$$w_{i,t} \approx \underbrace{\frac{1}{\tilde{\gamma}}\left(\frac{\mu_{i,t} - \beta'_{-i,t}\mu_{-i,t}}{\sigma^2_{i,\text{unspanned},t}}\right)}_{\text{main component}} + \underbrace{\frac{\theta}{\tilde{\gamma}\psi}\cdot\left(\frac{\sigma_{i,c-w,t} - \sigma_{-i,c-w,t}}{\sigma^2_{i,\text{unspanned},t}}\right)}_{\text{consumption-hedging component}}, \tag{A.4}$$

where $\sigma_{i,c-w,t}$ and $\sigma_{-i,c-w,t}$ are the covariances between $y_{i,t}$ and $y_{-i,t}$ with the log consumption-to-wealth ratio, respectively. We continue to use the notation of $\mu_{i,t}$, $\beta_{-i,t}$, and $\sigma^2_{i,\text{unspanned},t}$ from Equation (10). They are now based on y_t, rather than r_t, but the difference is quantitatively minor. Plugging in Equation (A.4) into Equation (9) leads to the following modified demand elasticity:

$$\begin{aligned}
\eta_{i,t} &= 1 + \frac{1}{w_{i,t}}\left(\frac{1}{\tilde{\gamma}\sigma^2_{i,\text{unspanned},t}}\right)\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right) + \frac{1}{w_{i,t}\tilde{\gamma}\sigma^2_{i,\text{unspanned},t}}\frac{\theta}{\psi}\frac{\partial \sigma_{i,c-w,t}}{\partial \log(P_{i,t})} \\
&= 1 + \underbrace{\frac{1}{\mu_{i,t,\text{unspanned}}}\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right)}_{\text{main component}} + \underbrace{\frac{\theta}{\psi\mu_{i,t,\text{unspanned}}}\cdot\frac{\partial \sigma_{i,c-w,t}}{\partial \log(P_{i,t})}}_{\text{consumption-hedging component}}
\end{aligned}$$

□

Proof of Proposition 4. From Equation (7), the unspanned return is $\mu_{i,\text{unspanned}} = \mu_i - \beta'_{-i}\mu_{-i}$. Under a factor model, we can write the expected return on asset i as:

$$\mu_i = \alpha_i^F + (\beta_i^F)' \, \mathbb{E}[f],$$

where F is the expected return of the factors. Similarly, we can write:

$$\mu_{-i} = \alpha_{-i}^F + (\beta_{-i}^F)' \, \mathbb{E}[f]$$

Plugging this into the the expression for $\mu_{i,\text{unspanned}}$, we can write the unspanned return as:

$$\alpha_i = \underbrace{\alpha_i^F - \beta'_{-i}\alpha_{-i}^F}_{\equiv \Delta\overline{\alpha}_i} + \underbrace{(\beta_i^F)'\mathbb{E}[f] - \beta'_{-i}(\beta_{-i}^F)'E[f]}_{\equiv \Delta\overline{\beta}_i}$$

To prove this, then it should be noted that:

$$\begin{aligned} \beta_i^F &= \mathbb{V}\text{ar}^{-1}(f) \, \text{Cov}(f, r_i) \\ &= \mathbb{V}\text{ar}^{-1}(f) \, \text{Cov}(W'r, \iota_i' r) \\ &= \mathbb{V}\text{ar}^{-1}(f) W' \Sigma \iota_i \\ &= \mathbb{V}\text{ar}^{-1}(f) W' \Sigma_i \end{aligned}$$

where ι_i is a vector of zeros, but 1 in the i^{th} spot.

We can also calculate

$$\begin{aligned} \beta_{-i} &= \mathbb{V}\text{ar}^{-1}(r_{-i})\text{Cov}(r_{-i}, r_i) = \mathbb{V}\text{ar}^{-1}(r_{-i})\text{Cov}(I_{-i}r, \iota_i' r) \\ &= \Sigma_{-i,-i}^{-1} I_{-i} \Sigma \iota_i = \Sigma_{-i,-i}^{-1} I_{-i} \Sigma_i \\ \beta_{-i}^F &= \mathbb{V}\text{ar}^{-1}(f)\text{Cov}(f, r_{-i}) = \mathbb{V}\text{ar}^{-1}(f)\text{Cov}(W'r, I_{-i}r) \\ &= \mathbb{V}\text{ar}^{-1}(f) W' \Sigma I'_{-i} \end{aligned}$$

Then putting this together we have:

$$\beta'_{-i}(\beta_{-i}^F)'\mathbb{E}[f] = \Sigma_i' I'_{-i} \Sigma_{-i,-i}^{-1} I_{-i} \Sigma W \mathbb{V}\text{ar}^{-1}(f) \mathbb{E}[f]$$

Thus if we define:

$$W^*_{-i} = \underbrace{I'_{-i}\Sigma^{-1}_{-i,-i}I_{-i}\Sigma}_{\text{projection-like matrix}} \; W,$$

we have:

$$\mu_{i,\text{unspanned}} = \underbrace{\alpha_i - \beta'_{-i}\alpha_{-i}}_{\alpha_{i,\text{unspanned}}} + \underbrace{\Sigma'_i(W - W^*_{-i})\omega^f}_{\text{weight dependence on } i}$$

It should be noted that pre-multiplying by I'_{-i} changes the $N-1$ dimensional vector to be N dimensional, with all terms unchanged but zero in the i^{th} spot. So the projection-like matrix $I'_{-i}\Sigma^{-1}_{-i,-i}I_{-i}\Sigma$ is close to the identity matrix if dropping the asset does not change the covariance matrix very much. Another way to see this is to note that:

$$I'_{-i}\Sigma^{-1}_{-i,-i}I_{-i}\Sigma = I'_{-i}\left(I_{-i}\Sigma I'_{-i}\right)^{-1} I_{-i}\Sigma.$$

□

A.2 Wealth effects are small

Our derivations in Section 2 ignore wealth effects. This section shows that the effect of wealth changes on asset-level demand elasticity is minimal, as the effect of price changes in an individual asset has limited effect on overall wealth for reasonably diversified portfolios.

Let A_t represent the assets under management (AUM) for a fund or the wealth for an investor. Recall that the quantity of shares demanded, $Q_{i,t}$, is given by $Q_{i,t} = \frac{A_t w_{i,t}}{P_{i,t}}$. The demand elasticity that incorporates wealth effects is:

$$\eta_{i,t} \equiv -\frac{\partial \log(Q_{i,t})}{\partial \log(P_{i,t})} = \underbrace{1 - \frac{\partial \log(w_{i,t})}{\partial \log(P_{i,t})}}_{\text{original term}} - \underbrace{\frac{\partial \log(A_t)}{\partial \log(P_{i,t})}}_{\text{wealth effect}}.$$

Focusing on the wealth effect, A_t can be expressed as:

$$A_t = A_{t-1}\left[w'_{t-1}(r_t + R_{f,t}\iota) + (1 - \iota' w_{t-1})R_{f,t}\right],$$

where r_t represents the vector of excess returns, $R_{f,t}$ the gross risk-free rate, and ι a vector of ones. In response to the change of the price of asset i, the wealth effect on demand elasticity is given by:

$$-\frac{\partial \log(A_t)}{\partial \log(P_{i,t})} = -\left(\frac{A_{t-1}}{A_t}\right)\left(\frac{P_{i,t}}{P_{i,t-1}}\right) w_{i,t-1}.$$

Typically, both AUM and stock prices exhibit only minor fluctuations, keeping the first two terms close to one. Additionally, $w_{i,t-1}$ is small in a well-diversified portfolio. Thus, wealth effects on demand elasticities are negligible.

A.3 Understanding low price pass-through

Section 3.2 found price pass-through to be significantly lower than one in empirical estimates. How should we understand this result? We first provide intuition that low price pass-through means that the variation of expected returns is persistent. We then note that many dynamic equilibrium models feature low price pass-throughs.

Persistent variation in expected returns. For the simplest example with persistent expected return variation, consider the Gordon growth model:

$$P = \frac{D}{r - g}$$

where D is the current dividend, r is the expected return, and g is the dividend growth rate. Consider a cash flow-unrelated price movement driven by a permanent change in r. Then, the implied price pass-through is simply the dividend yield,

$$-\frac{dr}{d\log(P)} = \frac{D}{P}.$$

To quantify this, note that after World War II, the average annual dividend yield of the U.S. stock market is 0.032. At a monthly frequency, this translates to price pass-through of approximately

$0.032/12 = 0.0027$, which is a fraction of the point estimate of 0.014 in Section 3.2. While this is an extreme example with fully persistent expected return variation, many important stock return predictors that explain a large amount of price variation—size, value, etc.—are associated with slowly-varying expected return variation (van Binsbergen et al., 2023). As a consequence, they are all associated with low price pass-throughs, a finding that we empirically confirm in Appendix B.1.

Equilibrium models with low price pass-through. Because our goal is to study demand elasticity as a portfolio choice problem, in this paper, we take expected return variation as given and do not take a stance on its origin. However, one may still wonder whether low price pass-throughs can be sustained in equilibrium. The answer is yes: many dynamic models explicitly predict low price pass-throughs. In fact, whenever expected return variation is induced by slow-moving preferences changes and trading flows, price pass-through is naturally low. Such models are featured in Duffie (2010), Kozak, Nagel, and Santosh (2018), Gabaix and Koijen (2022), among others.

B Empirical Details

B.1 Price pass-through for different anomalies

In Section 3.2, we estimated price pass-through using the van Binsbergen et al. (2023) "price wedge", which combines the effect of 57 stock characteristics. However, some researchers highlight that characteristics in different categories may have different properties. In this section, we decompose the price wedge by characteristic categories and separately estimate price pass-throughs for each.

Following Freyberger et al. (2020), we group characteristics into six categories: (1) "past returns"-based predictors (e.g., momentum and short-term reversal); (2) "investment"-related characteristics (e.g., the annual percentage change in total assets or the change in PP&E and inventory over total assets); (3) "profitability"-related characteristics (e.g., gross profitability over the book-value of equity or return on operating assets); (4) "intangibles" (e.g., operating accruals or tangibility); (5) "value"-related characteristics (e.g., the book-to-market ratio or cash to total assets); and (6) "trading frictions" (e.g., the average daily bid-ask spread and standard deviation of daily volume). See Table 1 in Freyberger et al. (2020) for the complete list of characteristics in each group.

We begin by computing the first principal component (PC1) of the characteristics in each of the six categories described above, treating each (stock, year-month) as an independent observation. Next, we perform linear regressions on the full sample to project the price wedge from van Binsbergen et al. (2023) for each stock i onto the PC1 of each of the six categories:

$$\log\left(\frac{P_{i,t}}{\tilde{P}_{i,t}}\right) = a_j + b_j \times \text{PC1}_{i,t}^j + \epsilon_{i,t}^j, \quad j = 1, \ldots, 6, \tag{B.1}$$

where $\text{PC1}_{i,t}^j$ represents the first PC of characteristics of stock i in category j in month t. Finally, we run Fama-MacBeth regressions at different horizons, similar to specification (12), with the independent variables replaced by the fitted values from regression (B.1).

The estimated (negative of) price pass-through for horizons $H = 1$, 3, 6, and 12 months are reported in Table B.1. At the one-month horizon, all anomalies other than the ones in the trading friction category have low price pass-throughs. For the categories of past returns, investment, and value, price pass-throughs are not statistically distinguishable from zero. For profitability and intangibles, the price pass-throughs are actually negative, indicating that price movements in those components exhibit continued momentum. The only category with large price pass-through is trading frictions, but it explains only a small fraction of overall price variation. Specifically, column (5) reports the R-squared when projecting price wedge on the PC1 of each category of characteristics. While value and intangibles have high R-squared, implying that these characteristics explain a higher fraction of price wedge variation, the other categories have low R-squared. Trading frictions, in particular, only explain 2.7% of overall price wedge variations.

B.2 Calibrating demand elasticity with heterogeneous stocks

Section 2.3 shows that classical asset pricing models predict very high demand elasticities based on a calibration that assumes stocks to be symmetric. In this section, we use simulations to introduce heterogeneity in asset expected returns, volatilties, and correlations, and see how this affects the demand elasticity calibration.

In each simulation with 1,000 stocks, to calibrate the term $\iota' \Sigma_t^{-1} \mu_t$ in Equation (11), we draw μ_t as a vector of random normal variables with mean of $0.06/12$ per month (annual return of 6%) and a standard deviation of $0.01/12$. For the covariance matrix, we draw $G \sim \mathcal{W}(V, 2N)$, where $\mathcal{W}(V, d)$ is the Wishart distribution with scale matrix V and degrees of freedom d. Matrix V is

Independent variable	Estimated coefficient β_H $H = 1$	3	6	12	R^2	obs
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\text{PC1}}^{\text{past returns}}$	−0.002	0.063	0.164	0.438	0.0405	526,807
	(0.044)	(0.163)	(0.424)	(0.813)		
$\widehat{\text{PC1}}^{\text{investment}}$	0.034	0.094	0.138	0.240	0.0112	526,807
	(0.023)	(0.099)	(0.188)	(0.232)		
$\widehat{\text{PC1}}^{\text{profitability}}$	4.566***	13.635**	28.195**	51.058*	0.0002	526,807
	(1.644)	(6.194)	(13.494)	(28.616)		
$\widehat{\text{PC1}}^{\text{intangibles}}$	0.037**	0.107*	0.207	0.432	0.1013	526,807
	(0.016)	(0.062)	(0.176)	(0.369)		
$\widehat{\text{PC1}}^{\text{value}}$	0.022	0.061	0.122	0.179	0.2202	526,807
	(0.016)	(0.086)	(0.265)	(0.626)		
$\widehat{\text{PC1}}^{\text{trading frictions}}$	−0.235***	−0.645***	−1.206**	−2.208**	0.0270	526,807
	(0.061)	(0.251)	(0.538)	(0.972)		
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.1. Estimating price pass-through for characteristics groups.
We estimate price pass-through for each of the six characteristics groups in Freyberger et al. (2020) using Fama-MacBeth regressions:

$$r_{i,t+1\to t+H} = \alpha_H + \beta_H \cdot \log \widehat{\left(P_{i,t}/\tilde{P}_{i,t}\right)}^{j}_{i,t} + \epsilon^{j}_{i,t+1\to t+H}, \quad j = 1,\ldots,6,$$

where the dependent variable is the log return of stocks in months $t+1$ to $t+H$. The independent variables are the fitted values of the "price wedge" measure in van Binsbergen et al. (2023) projected onto the first PC of characteristics of stock i in category j in month t. Columns 1 through 4 report estimated regression coefficients β_H for horizons $H = 1$, 3, 6, and 12 months. Column (5) reports the R-squared from regressing price wedge on PC1s of the characteristics from regression (B.1). Column (6) reports the number of stock-months used in each specification. The standard errors of the Fama-MacBeth forecasting coefficients are estimated using the Newey-West procedure with the number of lags equal to the forecasting horizon H.

constructed so that all pairwise correlations are 0.3 and the stock volatilities are 10%. We then set $\Sigma_t = G/N$, which in expectation equals V. While V assumes uniform pairwise covariances, Σ_t introduces variation, making the simulations realistic. The degrees of freedom $d = 2N$ ensures that the matrix is well-conditioned.[1] We do 1,000 simulations, yielding $\iota'\Sigma_t^{-1}\mu_t \approx 3.35$ on average. The average R^2 from regression (6) becomes 65%, indicating that for an asset with total volatility of 10%, the unspanned volatility is reduced to 6% compared to 8.4% in Section 2.3. Plugging these results into Equation (11) again yields unspanned returns of approximately 0.0012% as in the main calibration.

[1]We require at least $d \geq N$, but larger values of d help ensure that the matrix is well-conditioned

B.3 Quantifying the volatility and correlation components

In this section, we quantify the "volatility" and "correlation" components of demand elasticity in Proposition 2.

The volatility component. For this, we need to estimate the unspanned variance of stocks ($\sigma^2_{i,\text{unspanned},t}$) and how the stock return variance $\sigma^2_{i,t}$ responds to price movements.

For the first term, we obtain the daily return residuals for each stock-year after controlling for the Fama-French five factors, and then compute its average monthly variance to be 0.027 in our sample. For the second term, we regress the monthly return variance (computed from daily returns) for each stock i and month t in a panel regression:

$$\sigma^2_{i,t} = b_0 + b_1 r_{i,t-1} + b_2 \sigma^2_{i,t-1} + \epsilon_{i,t},$$

where $r_{i,t-1}$ is the lagged monthly log stock return. We cluster standard errors by year-month and stock, and the results are reported in Table B.2. Consistent with the "leverage effect" of Black (1976), stock return variance co-moves negatively with returns. Over the full sample, when controlling for both stock and year-month fixed effects, we estimate the response of monthly return variance to log returns to be around -0.065. Results in columns (4) through (9) show that the relationship is relatively stable over subsamples.

The correlation component. Davis (2024) considers an $N \times K$ matrix of K characteristics for stocks in time t and estimates a model where the covariance matrix and vector of expected returns are linear functions of predictors:

$$\mu_t = Z_t \Theta_\mu, \quad \Gamma_t = Z_t \Theta_\Gamma, \quad \Sigma_t = \Gamma_t \Gamma_t' + \Theta_\zeta I, \tag{B.2}$$

where Θ_μ and Θ_Γ are vectors of parameters that govern the mean and covariance matrix, respectively, and $\Theta_\zeta > 0$ is a positive scalar that controls unspanned variance. Some characteristics are functions of prices (e.g. book-to-market), and therefore a movement in asset price changes both the mean returns μ_t and the covariance matrix through Γ_t. In Davis (2024), the model in Equation (B.2) is estimated using maximum likelihood with monthly return data and normalized versions of the

	Dependent variable: $\sigma^2_{i,t}$								
	Full sample			1970 - 1994			1995 - 2019		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$r_{i,t-1}$	-0.077^{***}	-0.063^{***}	-0.065^{***}	-0.070^{***}	-0.058^{***}	-0.064^{***}	-0.081^{***}	-0.062^{***}	-0.059^{***}
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)
$\sigma^2_{i,t-1}$	0.633^{***}	0.472^{***}	0.456^{***}	0.689^{***}	0.537^{***}	0.528^{***}	0.596^{***}	0.392^{***}	0.366^{***}
	(0.010)	(0.011)	(0.011)	(0.013)	(0.014)	(0.014)	(0.013)	(0.012)	(0.011)
Stock FE	N	Y	Y	N	Y	Y	N	Y	Y
Time FE	N	N	Y	N	N	Y	N	N	Y
Obs	2,136,330	2,136,330	2,136,330	960,031	960,031	960,031	1,176,299	1,176,299	1,176,299
R^2	37.51%	43.55%	44.59%	44.75%	49.70%	50.36%	33.20%	41.55%	42.98%
Note:									*p<0.1; **p<0.05; ***p<0.01

Table B.2. Estimating the effect of prices on return variance.
We estimate the Black (1976) effect using panel regressions of monthly stock return variance the lagged monthly log stock return. Stock return variance is estimated using daily Fama-French 5-factor model residuals. Columns (1) through (3) uses the full sample. Columns (4) through (6) use the first half of the sample. Columns (7) through (9) use the second half of the sample. When using the same sample, the regressions differ in whether they include stock- or time-fixed effects. Standard errors are clustered by month and stock.

characteristics from Freyberger et al. (2020). The implied demand elasticity due to price-induced changes in covariance is approximately 1.4.

B.4 Impact of consumption hedging: details

To assess the quantitative importance of the consumption-hedging component in Equation (15), we perform a calibration using data. Specifically, we estimate the sensitivity of the covariance between stock returns and the log consumption-to-wealth ratio, $\sigma_{i,c-w,t}$, to changes in stock prices. Our approach involves three steps:

Step 1: Estimating unexpected returns. First, we estimate the unexpected component of stock returns, $r^{\text{unex}}_{i,t+1}$, by regressing realized quarterly returns on the (log) price wedge instrument from van Binsbergen et al. (2023), controlling for stock fixed effects:

$$r_{i,t+1} = \alpha_i + \beta_1 \, PW_{i,t} + \varepsilon_{i,t+1}, \tag{B.3}$$

where $r_{i,t+1}$ is the realized return for stock i in quarter $t+1$, $PW_{i,t} = \log\left(P_{i,t}/\tilde{P}_{i,t}\right)$ is the (log) price wedge from van Binsbergen et al. (2023), and α_i captures stock fixed effects. The residuals from this regression, $\varepsilon_{i,t+1}$, represent the unexpected returns: $r^{\text{unex}}_{i,t+1} = \varepsilon_{i,t+1}$.

Step 2: Estimating the sensitivity of the consumption-hedging term. Next, we estimate the sensitivity of the covariance between unexpected returns and the log consumption-to-wealth ratio deviations, cay_{t+1}, with respect to the price wedge. We achieve this by regressing the product of unexpected returns and cay_{t+1} on the (log) price wedge, again controlling for stock fixed effects:

$$r^{\text{unex}}_{i,t+1} \cdot cay_{t+1} = \alpha_i + \beta_2\, PW_{i,t} + \eta_{i,t+1}, \tag{B.4}$$

where cay_{t+1} is the consumption-to-wealth ratio log deviations from Lettau and Ludvigson (2001). Since the cay data is available quarterly, we use quarterly returns. We obtain the cay data from Amit Goyal's website (https://sites.google.com/view/agoyal145), as the data is frequently updated there. In regression (B.4), β_2 captures the sensitivity of the consumption-hedging term to (cash flow-unrelated) changes in the asset price $\left(\frac{\partial \sigma_{i,c-w,t}}{\partial \log(P_{i,t})}\right)$. Note that the expected value of the left-hand side is $\mathbb{E}_t[r^{\text{unex}}_{i,t+1} \cdot cay_{t+1}] = \sigma_{i,c-w,t}$.

Results. Table B.3 presents the results of the regressions discussed above. In column (1), we report the estimates from Equation (B.3). The negative and significant coefficient β_1 indicates that higher price wedges are associated with lower expected returns, as expected.[2] In column (2), we present the estimates from Equation (B.4). The coefficient β_2 is small in magnitude and not statistically significant, suggesting that the sensitivity of the consumption-hedging term to price changes is negligible.

When we convert the quarterly estimate of $\beta_2 = -0.0005$ to a monthly value (by dividing by 3), we get approximately -0.00017. This value is two orders of magnitude smaller than our pass-through estimate of 0.014, indicating that the consumption-hedging component has a minimal impact on overall demand elasticity.

[2]It is worth noting that this coefficient should not be interpreted as the quarterly price pass-through, as we also control for stock fixed effects. By doing so, we incorporate information from the future, and thus this is not a true forecasting regression.

	Dependent variable:	
	r_{t+1}	$r^{unex}_{t+1} cay_{t+1}$
	(1)	(2)
(log) price wedge	-0.1918***	-0.0005
	(0.0167)	(0.0004)
Observations	425,777	425,777
R^2	0.0104	0.0001
Stocked fixed effects	Yes	Yes
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table B.3. Consumption hedging term price sensitivity regressions
This table presents the regression estimates used to calibrate the importance of the Epstein-Zin intertemporal hedging term for demand elasticity. First, we regress stock-level quarterly returns on the (log) price wedge from van Binsbergen et al. (2023) at time t, with stock fixed effects, as shown in column (1). We then take the residual for this regression, labeled r^{unex}_{t+1}. Next, we regress $r^{unex}_{t+1} cay_{t+1}$ on the (log) price wedge with stock fixed effects, which is a quarterly measure of $\partial \sigma_{i,c-w,t} / \partial \log(P_{i,t})$ from Equation (15), as shown in column (2).

B.5 Impact of transaction costs: details

To study the impact transaction costs on portfolio choice, we use a slightly adapted version of the model in Gârleanu and Pedersen (2013) . In their model, the cost-optimized portfolio weight vector, w^*_t, is a linear combination of the existing portfolio and the "aim" portfolio:

$$w^*_t = (1 - s_{\text{aim}}) \; w_{t \leftarrow t-1} + (s_{\text{aim}}) \; \text{aim}_t,$$

where $w_{t \leftarrow t-1}$ represents the previous period's portfolio weight, passively adjusted based on price movements in the current period. The "aim" portfolio is a weighted average of current and expected future optimal portfolios in the absence of transaction costs:

$$\text{aim}_t = \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}}) \, (\rho_{\text{aim}})^{\tau} \, \mathbb{E}_t [w_{t+\tau}], \tag{B.5}$$

where w_t are the optimal portfolio weights without transaction costs. The parameters s_{aim} and ρ_{aim} are scalar values, each bounded between zero and one.

Let $\eta_{i,t\to t+\tau}$ be the elasticity of the future not-cost-optimized portfolio:

$$\eta_{i,t\to t+\tau} = 1 - \frac{1}{\mathbb{E}_t\left[w_{t+\tau}\right]}\frac{\partial \mathbb{E}_t\left[w_{t+\tau}\right]}{\partial \log(P_{i,t})}.$$

In terms of elasticity, assuming $w_{i,t\leftarrow t-1} > 0$ and $\mathbb{E}_t\left[w_{i,t+\tau}\right] > 0$ for all τ, the cost-optimized elasticity, $\eta^*_{i,t}$, can be expressed as:

$$\begin{aligned}
\eta^*_{i,t} &\equiv 1 - \frac{\partial \log(w^*_{i,t})}{\partial \log(P_{i,t})} \\
&= 1 - \frac{1}{w^*_{i,t}}\frac{\partial w^*_{i,t}}{\partial \log(P_{i,t})} \\
&= 1 - \frac{1}{w^*_{i,t}}\left((1 - s_{\text{aim}})\,\frac{\partial w_{t\leftarrow t-1}}{\partial \log(P_{i,t})} + (s_{\text{aim}})\,\sum_{\tau=0}^{\infty}(1-\rho_{\text{aim}})\,(\rho_{\text{aim}})^{\tau}\,\frac{\partial \mathbb{E}_t\left[w_{t+\tau}\right]}{\partial \log(P_{i,t})}\right).
\end{aligned} \tag{B.6}$$

Note that by definition of passively floated portfolio weights, we must have:

$$\frac{\partial \log(w_{t\leftarrow t-1})}{\partial \log(P_{i,t})} = 1. \tag{B.7}$$

In other words, when prices move up 1%, passive weights also increase by 1%. This makes the passively floated component of the portfolio completely inelastic (i.e., elasticity is $0 = 1 - 1$). Equation (B.7) implies:

$$\frac{\partial w_{t\leftarrow t-1}}{\partial \log(P_{i,t})} = w_{t\leftarrow t-1}.$$

We can substitute this into Equation (B.6) above and proceed with calculating the cost-optimized elasticity:

$$
\begin{aligned}
\eta^*_{i,t} &= 1 - \frac{1}{w^*_{i,t}} \left((1 - s_{\text{aim}})\, w_{t \leftarrow t-1} + (s_{\text{aim}}) \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \frac{\partial \mathbb{E}_t [w_{t+\tau}]}{\partial \log(P_{i,t})} \right) \\
&= 1 - \frac{1}{w^*_{i,t}} \left((1 - s_{\text{aim}})\, w_{t \leftarrow t-1} - (s_{\text{aim}}) \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \mathbb{E}_t [w_{t+\tau}] \left(\eta_{i,t \to t+\tau} - 1 \right) \right) \\
&= 1 - \frac{1}{w^*_{i,t}} \left((1 - s_{\text{aim}})\, w_{t \leftarrow t-1} + (s_{\text{aim}}) \text{aim}_{i,t} - (s_{\text{aim}}) \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \mathbb{E}_t [w_{t+\tau}] \left(\eta_{i,t \to t+\tau} \right) \right) \\
&= 1 - \frac{1}{w^*_{i,t}} \left(w^*_{i,t} - (s_{\text{aim}}) \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \mathbb{E}_t [w_{t+\tau}] \left(\eta_{i,t \to t+\tau} \right) \right) \\
&= \frac{s_{\text{aim}}}{w^*_{i,t}} \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \mathbb{E}_t [w_{t+\tau}] \left(\eta_{i,t \to t+\tau} \right).
\end{aligned}
$$

Note that $\eta_{i,t \to t+\tau}$ should decay as τ increases because portfolio weights in the more distant future should be less sensitive to current prices. To further simplify, we assume a geometric decay of $\eta_{i,t \to t+\tau}$:

$$
\eta_{i,t \to t+\tau} = \rho_{\eta}^{\tau}\, \eta_{i,t}, \tag{B.8}
$$

where $\rho_{\eta} \in (0,1)$ is the decay coefficient. If $\rho_{\eta} \approx 1$, then using the definition of the aim portfolio in Equation (B.5), we can write:

$$
\begin{aligned}
&\sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \mathbb{E}_t [w_{t+\tau}] \rho_{\eta}^{\tau}\, \eta_{i,t} \\
&\approx \text{aim}_{i,t} \sum_{\tau=0}^{\infty} (1 - \rho_{\text{aim}})\, (\rho_{\text{aim}})^{\tau}\, \rho_{\eta}^{\tau}\, \eta_{i,t} \\
&= \text{aim}_{i,t} \left(\frac{1 - \rho_{\text{aim}}}{1 - \rho_{\eta} \rho_{\text{aim}}} \right) \eta_{i,t}
\end{aligned}
$$

Thus, we arrive at the approximation given in Equation (16) of the paper:

$$
\eta^*_{i,t} \approx s_{\text{aim}} \left(\frac{\text{aim}_{i,t}}{w^*_{i,t}} \right) \left(\frac{1 - \rho_{\text{aim}}}{1 - \rho_{\eta} \rho_{\text{aim}}} \right) \eta_{i,t}. \tag{B.9}
$$

Equation (B.9) shows that, depending on the two parameters s_{aim} and ρ_{aim}, a cost-optimized portfolio can result in very inelastic demand. As Gârleanu and Pedersen (2013) discuss, if transaction costs are high, s_{aim} is close to 0. In this case, the elasticity becomes nearly zero. If ρ_{aim} is close to one and $\rho_\eta < 1$ then the elasticity is further dropped through the aim portfolio being less sensitive to prices today. This logic is intuitive: the aim portfolio is a long-run portfolio when ρ_{aim} is large, so if $\rho_\eta < 1$, then this long-run portfolio is less sensitive to price variation today.

Using a method similar to that of Gârleanu and Pedersen (2013), we calibrate s_{aim} to be approximately 0.03 and ρ_{aim} to be approximately 0.97. Specifically, we use an annual discount rate of 4%, and, following Gârleanu and Pedersen (2013), we set the absolute risk aversion parameter to $\gamma = 10^{-9}$. For the trading cost parameter λ, we adopt their more conservative estimate of $\lambda = 10 \times 10^{-7}$. This implies $s_{\text{aim}} = a/\lambda \approx 0.03$ and $\rho_{\text{aim}} = 1 - \gamma/(\gamma + a) \approx 0.97$, where parameter a can be calculated from Equation (9) in Gârleanu and Pedersen (2013).

These values suggest large transaction costs. For robustness, we also consider values that imply lower transaction costs: $s_{\text{aim}} = 0.1$ and $\rho_{\text{aim}} = 0.9$. To calibrate ρ_η, we estimate an AR(1) regression of the log-price wedge from van Binsbergen et al. (2023), which yields an estimated ρ_η of 0.99. It is worth noting that if price pass-throughs are generally low, then price variation must also be persistently high on average.

B.6 Computing unspanned returns in factor models

This section explains how we compute unspanned returns for different factor models in Section 5.2. Proposition 5 shows a computationally efficient formula. Specifically, to compute unspanned returns, we need four components: the unspanned alpha, unspanned variance, asset portfolio weights across factors, and the MV portfolio weight across factors (ω^f). Unspanned alpha calculation is discussed in Appendix C.4. Unspanned variances can simply be calculated by taking the inverse of each term in the diagonal of Σ_t^{-1}. The other terms in Proposition 5 have already been calculated following the methods described in Section 5.2.[3]

[3]Note that the $W_i\omega^f$ term is invariant to scale. In other words, consider $\tilde{W}$ defined as $\tilde{W} = cW$, where c is a scalar. Then $\tilde{W}_i = cW_i$. However, $\tilde{W}_i\tilde{\omega}^f = cW_i \text{Var}^{-2}(cf)\mathbb{E}[cf] = W_i\omega^f$. Thus, scaling the factor model weights up and down has no effect on demand elasticity, which is sensible, as elasticity is defined in percentage changes.

Proposition 5. *We can write the unspanned return of asset i as:*

$$\mu_{i,unspanned} = \alpha_{i,unspanned} + \underbrace{\sigma^2_{i,unspanned}}_{\text{unspanned variance}} \times \underbrace{W_i \omega^f}_{\text{weight of asset } i \text{ in SDF portfolio}},$$

where $\sigma^2_{i,unspanned} = \sigma_i^2 - \beta'_{-i}\Sigma_{-i,-i}\beta_{-i}$ is the return variance of asset i not spanned by other assets.

Proof. Here we prove the result for asset 1, and this obviously generalizes to asset i. From Equation (18), we have:

$$W^*_{-1} = \left(I'_{-1}\Sigma^{-1}_{-1,-1}I_{-1}\Sigma\right)W$$

Note that

$$I'_{-1}\Sigma^{-1}_{-1,-1}I_{-1} = \begin{bmatrix} 0 & \mathbf{0}'_{N-1} \\ \mathbf{0}_{N-1} & \Sigma^{-1}_{-1,-1} \end{bmatrix},$$

where $\mathbf{0}_{N-1}$ is an $(N-1)$ dimensional column vector of zeros. We can also write:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \Sigma'_{-1} \\ \Sigma_{-1} & \Sigma_{-1,-1} \end{bmatrix}.$$

Thus we can calculate:

$$I'_{-1}\Sigma^{-1}_{-1,-1}I_{-1}\Sigma = \begin{bmatrix} 0 & \mathbf{0}'_{N-1} \\ \Sigma_{-1,-1}\Sigma_{-1} & I \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}'_{N-1} \\ \beta_{-1} & I \end{bmatrix}.$$

We partition W into the following:

$$W = \begin{bmatrix} W_1 \\ W_{-1} \end{bmatrix},$$

where W_1 is the row vector containing the i^{th} row of W. Then we can write:

$$W^*_{-1} = \left(I'_{-1}\Sigma^{-1}_{-1,-1}I_{-1}\Sigma\right)W = \begin{bmatrix} 0 & \mathbf{0}'_{N-1} \\ \beta_{-1} & I \end{bmatrix}\begin{bmatrix} W_1 \\ W_{-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \beta_{-1}W_1 + W_{-1} \end{bmatrix}.$$

Thus we have the following:

$$W - W_{-1}^{*} = \begin{bmatrix} W_1 \\ -\beta_{-1} W_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -\beta_{-1} \end{bmatrix} W_1.$$

Furthermore:

$$\Sigma_1' \begin{bmatrix} 1 \\ -\beta_{-1} \end{bmatrix} = \sigma_1^2 - \Sigma_{-1}\beta_{-1} = \sigma_1^2 - \beta_{-1}\Sigma_{-1,-1}\beta_{-1} = \sigma_{1,\text{unspanned}}^2.$$

So putting this altogether means that:

$$\Sigma_1'(W - W_{-1}^{*})\omega^f = \sigma_{1,\text{unspanned}}^2 W_1 \omega^f.$$

$\square$

C Additional Considerations

This section discusses various additional considerations in MV portfolio choice. Sections C.1 and C.2 discuss the case of negative and small portfolio weights. Section C.3 explains why extending investment horizon to quarterly or annual does not materially affected predicted demand elasticities. Section C.4 examines extensions of the PCA factor models with non-zero alphas.

C.1 Negative portfolio weights

Demand elasticity is calculated using log quantities, and thus is undefined for short position. Due to this reason, we follow Koijen and Yogo (2019) to compute demand elasticities for stocks with positive portfolio weights in Section 3.4. In this section, we extend the definition of demand elasticity to short positions. Under this extended definition, demand elasticities for long and short positions are quantitatively similar, so our main insights also carry over to short positions.

We extend the definition of demand elasticity to

$$\eta_{i,t}^{\pm} = 1 - \frac{1}{|w_{i,t}|}\left(\frac{\partial w_{i,t}}{\partial \log(P_{i,t})}\right),$$

which is identical to the regular definition for long positions. For short positions, the $|w_{i,t}|$ term means the interpretation is maintained that a positive elasticity implies a downward sloping demand curve, except that $\eta^{\pm}_{i,t}$ is in terms of a percentage change of the absolute value of the position size instead of just the position size.

Our decomposition of demand elasticity into two components also extend naturally to short positions. It is not hard to follow the proof of Proposition 1 to show that

$$\eta^{\pm}_{i,t} = 1 + \frac{1}{|\mu_{i,\text{unspanned},t}|}\left(-\frac{\partial \mu_{i,t}}{\partial \log(P_{i,t})}\right).$$

Thus, as long as unspanned returns are close to being symmetric around zero, demand elasticities would be close to symmetric between long and short positions. Figure 2 shows this to be the case in empirical estimates. We also verified that unspanned returns tend to have a symmetric distribution around zero for all the PCA factor models in Section 5.2. Therefore, elasticities are also low for short positions under this generalized definition of demand elasticities.

C.2 Small portfolio weights

Figure 2 shows that there are stocks for which the unspanned return is positive but very small (i.e., where the two lines almost intersect), and by our decomposition, an MV investor would have a very high demand elasticity in these stocks. In fact, as the unspanned return approaches zero, an MV investor's demand elasticity approaches infinity (Equation 9). However, the MV investor also has vanishingly small positions in these stocks (Equation 10), and thus would exert very little influence on the aggregate demand elasticities. When using more economically relevant quantities, such as portfolio holdings-weighted demand elasticities, we again find demand elasticities to be low.

To see that small portfolio positions have small effects on the aggregate demand elasticity, consider the following derivation for an arbitrary asset. We omit the asset and time subscripts for notational simplicity. Suppose there are a total of $j = 1, \ldots, J$ investors, each of which demand Q_j shares of the asset. Define aggregate shares demand as $Q = \sum_j Q_j$. The aggregate demand elasticity is given by:

$$
\begin{aligned}
\eta^{\text{agg}} \equiv -\frac{\partial \log(Q)}{\partial \log(P)} &= -\frac{1}{Q}\frac{\partial \sum_j Q_j}{\partial \log(P)} \\
&= \sum_j \frac{Q_j}{Q} \cdot \frac{\partial \log(Q_j)}{\partial \log(P)} \\
&= \sum_j \frac{Q_j}{Q} \cdot \eta_j,
\end{aligned}
$$

where η_j is the demand elasticity of investor j. Therefore, aggregate demand elasticity is simply the holdings-weighted demand elasticities across investors, meaning that investors with small holdings have a limited impact on the aggregate demand elasticity. For simplicity, we have assumed that all investors have long positions, but this can be extended to include short positions using the method in Appendix C.1.

C.3 Alternative investment horizons

Our analysis in Section 3 assumes a monthly investment horizon. In practice, some investors adjust their portfolios more slowly and have longer holding horizons. Naturally, price pass-throughs are higher over longer horizons. This raises a question: would investors with longer horizons have significantly higher demand elasticities?

Not necessarily, as unspanned returns may also increase with the investment horizon. Our main demand elasticity formula in Equation (4), reproduced below, can be applied to any investment horizon H,

$$
\eta_{i,t\to t+H} \approx 1 + \underbrace{\frac{1}{\mu_{i,\text{unspanned},t\to t+H}}}_{\text{weight responsiveness}} \times \underbrace{\left(-\frac{\partial \mu_{i,t\to t+H}}{\partial \log(P_{i,t})}\right)}_{\text{price pass-through}}.
$$

While an investor with longer horizon—e.g., quarterly instead of monthly—would foresee higher price pass-through, they would also have higher unspanned returns. If these two factors scale similarly with the investment horizon, their effects may offset, leaving the demand elasticity prediction unchanged.

To investigate this empirically, the first row of Table C.4 reports the point estimates of price

		Investment horizon H (months)			
		1	3	6	12
		(1)	(2)	(3)	(4)
Original	Price pass-through	0.014	0.040	0.079	0.157
	Unspanned returns	0.23%	1.25%	2.02%	3.40%
Monthly	Price pass-through	0.014	0.013	0.013	0.013
	Unspanned returns	0.23%	0.42%	0.34%	0.28%
	Implied demand elasticity	6.9	4.2	4.9	5.6

Table C.4. Demand elasticity by investment horizon
This table computes the demand elasticity of mean-variance investors for investment horizons of 1, 3, 6, and 12 months, respectively. The first row reports the estimated price pass-through based on the first row in Table 2. The second row reports the average unspanned return of all positive portfolio positions using the methodology in Section 3.3. The next two rows report the equivalent values after converting to monthly frequency. The last row reports the implied demand elasticity.

pass-through for horizons of $H = 1$, 3, 6, and 12 months, taken from the estimates in Table 2. To estimate unspanned returns at different horizons, we re-estimate the Fama-MacBeth regression in Section 3.3 for each horizons, and report the implied average unspanned returns for positive portfolio positions in the second row of Table 2.

In the next two rows, we report the implied *monthly* price pass-through and unspanned returns by dividing by the horizon H. The results indicate that both price pass-through and unspanned returns scale approximately linearly with the horizon. As a consequence, the predicted demand elasticities, shown in the last row, remain roughly constant across horizons. This indicates that, at least for investment horizons within a year, demand elasticity predictions do not vary significantly with the horizon.

C.4 Factor models with non-zero alphas

When implementing the PCA-style factor models in Section 5.2, we did not allow for alphas that are not captured by the factors. In this section, we relax this restriction and estimate the N-dimensional vector of alphas (α_t). For the RPCA and IPCA models, we estimate alphas using the methods in their respective original papers. While Lettau and Pelger (2020) does not allow for stock-level alphas, we compute it using the same method as the RPCA model. We then compute

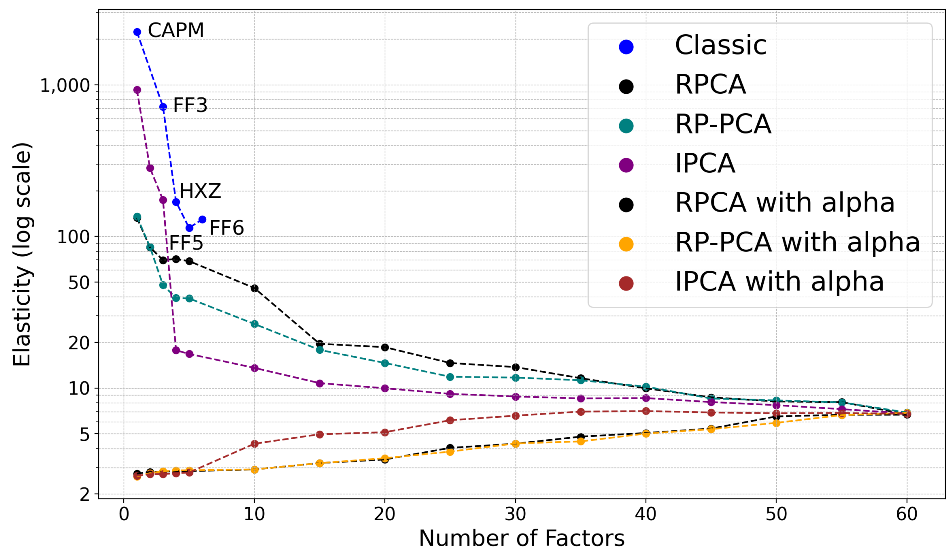

Figure C.1. Demand Elasticity Across Factor Models: Non-zero alphas
This figure is similar to Figure 3 but also adds PCA-style models with non-zero alphas. The number of factors is shown on the x-axis. The model-implied demand elasticity is plotted on a log scale on the y-axis, which uses price pass-through of 0.014. In the legend, "Classic" refers to traditional factor models. RPCA refers to the regression-PCA model in Chen et al. (2023). RP-PCA refers to the model in Lettau and Pelger (2020). IPCA refers to the model in Kelly et al. (2019).

unspanned alphas as $\alpha_{i,\text{unspanned},t} = \alpha_{i,t} - \beta'_{-i,t}\alpha_{-i,t}$ as in Equation (17). The betas are derived using the estimated covariance matrix Σ_t whose computation has been described in Section 5.2.

Figure C.1 and Table C.5 display the results. When comparing to the PCA-style models without alphas, including alphas further reduces elasticity. When the number of factors increase and converge towards the number of characteristics, the alphas are mechanically absorbed into factors and thus the difference shrinks. However, when there are fewer factors, the models exhibit lower demand elasticity when alphas are included. For instance, when there are 10 or fewer factors, the predicted demand elasticities are all below 5 for the three PCA-style models with alphas. This should be intuitive as unspanned alpha, by definition, is an asset-specific characteristic that is difficult to replicate, leading to even lower substitutability and lower demand elasticity.

Factor Model	Number of Factors	Unspanned Return	Weight Responsiveness	Demand Elasticity
		Classic		
CAPM	1	0.0006%	159678	2236.5
FF3	3	0.0020%	51145	717.0
HXZ	4	0.0083%	12018	169.3
FF5	5	0.0124%	8063	113.9
FF6	6	0.0109%	9164	129.3
		RPCA with alpha		
PCA factor model	5	0.7637%	131	2.8
PCA factor model	10	0.7340%	136	2.9
PCA factor model	15	0.6357%	157	3.2
PCA factor model	20	0.5875%	170	3.4
		RP-PCA with alpha		
PCA factor model	5	0.7484%	134	2.9
PCA factor model	10	0.7336%	136	2.9
PCA factor model	15	0.6363%	157	3.2
PCA factor model	20	0.5719%	175	3.4
		IPCA with alpha		
PCA factor model	5	0.7883%	127	2.8
PCA factor model	10	0.4249%	235	4.3
PCA factor model	15	0.3533%	283	5.0
PCA factor model	20	0.3401%	294	5.1

Table C.5. Demand Elasticity Across Factor Models: Non-zero alphas
This table presents the demand elasticity, calculated using Equation (9), for various factor models. The elasticity is computed with a pass-through of 0.014 and the average unspanned return of a factor model using Proposition 4. Note that Proposition 5 in the appendix offers a more efficient method for calculating the unspanned return of these models. This table essentially replicates Table 5, but includes alphas in the PCA factor models.

D Two-Asset Example

In this section, we use a two-asset example to illustrate why the degree of substitutability between assets matters for asset-level demand elasticity for a mean-variance investor. The following derivations also clarify where each term in Proposition 2 comes from.

Consider an investor with CARA utility who is forming an optimal portfolio at time t. In addition to a risk-free asset with an exogenously given risk-free rate, the investor can invest in two risky assets. The covariance matrix of the next-period dollar payoffs is defined as follows:

$$\begin{bmatrix} \Lambda_{1,t} & \Lambda_{1,2,t} \\ \Lambda_{1,2,t} & \Lambda_{2,t} \end{bmatrix} = \begin{bmatrix} \mathbb{V}\mathrm{ar}_t(P_{1,t+1} + D_{1,t+1}) & \mathrm{Cov}_t(P_{1,t+1} + D_{1,t+1}, P_{2,t+1} + D_{2,t+1}) \\ \mathrm{Cov}_t(P_{1,t+1} + D_{1,t+1}, P_{2,t+1} + D_{2,t+1}) & \mathbb{V}\mathrm{ar}_t(P_{2,t+1} + D_{2,t+1}) \end{bmatrix}$$

The investor's share demand is

$$Q_t = \frac{1}{\gamma} \begin{bmatrix} \Lambda_{1,t} & \Lambda_{1,2,t} \\ \Lambda_{1,2,t} & \Lambda_{2,t} \end{bmatrix}^{-1} \begin{bmatrix} \mathbb{E}_t[P_{1,t+1} + D_{1,t+1}] - R_{f,t}P_{1,t} \\ \mathbb{E}_t[P_{2,t+1} + D_{2,t+1}] - R_{f,t}P_{2,t}, \end{bmatrix}$$

where γ is the risk aversion parameter. We can explicitly expand out the matrix inverse above to get:

$$Q_{1,t} = \frac{\Lambda_{2,t}\left(\mathbb{E}_t[P_{1,t+1} + D_{1,t+1}] - R_{f,t}P_{1,t}\right) - \Lambda_{1,2,t}\left(\mathbb{E}_t[P_{2,t+1} + D_{2,t+1}] - R_{f,t}P_{2,t}\right)}{\gamma\left(\Lambda_{1,t}\Lambda_{2,t} - \Lambda_{1,2,t}^2\right)}$$
$$Q_{2,t} = \frac{\Lambda_{1,t}\left(\mathbb{E}_t[P_{2,t+1} + D_{2,t+1}] - R_{f,t}P_{2,t}\right) - \Lambda_{1,2,t}\left(\mathbb{E}_t[P_{1,t+1} + D_{1,t+1}] - R_{f,t}P_{1,t}\right)}{\gamma\left(\Lambda_{1,t}\Lambda_{2,t} - \Lambda_{1,2,t}^2\right)},$$

which can be slightly manipulated to get:

$$Q_{1,t} = \frac{\left(\mathbb{E}_t[P_{1,t+1} + D_{1,t+1}] - R_{f,t}P_{1,t}\right) - \frac{\Lambda_{1,2,t}}{\Lambda_{2,t}}\left(\mathbb{E}_t[P_{2,t+1} + D_{2,t+1}] - R_{f,t}P_{2,t}\right)}{\gamma\left(\Lambda_{1,t} - \frac{\Lambda_{1,2,t}^2}{\Lambda_{2,t}}\right)},$$
$$Q_{2,t} = \frac{\left(\mathbb{E}_t[P_{2,t+1} + D_{2,t+1}] - R_{f,t}P_{2,t}\right) - \frac{\Lambda_{1,2,t}}{\Lambda_{1,t}}\left(\mathbb{E}_t[P_{1,t+1} + D_{1,t+1}] - R_{f,t}P_{1,t}\right)}{\gamma\left(\Lambda_{2,t} - \frac{\Lambda_{1,2,t}^2}{\Lambda_{1,t}}\right)}.$$

Note that $\frac{\Lambda_{1,2,t}}{\Lambda_{2,t}}$ is the regression coefficient of asset 1 payoff on asset two. Therefore, the expressions above look like unspanned return in dollar terms divided by unspanned variance in dollar terms. To simplify notations, we define the following:

$$
\begin{aligned}
B_{-1,t} &= \frac{\Lambda_{1,2,t}}{\Lambda_{2,t}} \\
B_{-2,t} &= \frac{\Lambda_{1,2,t}}{\Lambda_{1,t}} \\
M_{1,t} &= \left(\mathbb{E}_t[P_{1,t+1} + D_{1,t+1}] - R_{f,t}P_{1,t}\right) \\
M_{2,t} &= \left(\mathbb{E}_t[P_{2,t+1} + D_{2,t+1}] - R_{f,t}P_{2,t}\right) \\
M_{1,\text{unspanned},t} &= M_{1,t} - B_{-1,t}M_{2,t} \\
M_{2,\text{unspanned},t} &= M_{2,t} - B_{-2,t}M_{1,t} \\
\Lambda_{1,\text{unspanned},t} &= \left(\Lambda_{1,t} - B_{-1,t}\Lambda_{1,2,t}\right) \\
\Lambda_{2,\text{unspanned},t} &= \Lambda_{2,t} - B_{-2,t}\Lambda_{1,2,t},
\end{aligned}
$$

where these unspanned returns and unspanned variance terms have the same interpretation as the $\mu_{i,\text{unspanned},t}$ and $\sigma^2_{i,\text{unspanned},t}$ terms, except that these are in dollar terms. Then, we can simplify the share demand expressions as:

$$
\begin{aligned}
Q_{1,t} &= \frac{M_{1,\text{unspanned},t}}{\gamma\Lambda_{1,\text{unspanned},t}} \\
Q_{2,t} &= \frac{M_{2,\text{unspanned},t}}{\gamma\Lambda_{2,\text{unspanned},t}}
\end{aligned}
$$

We can then directly calculate the elasticity for asset 1 as:

$$
\begin{aligned}
\eta_{i,t} &= -\frac{\partial \log(Q_{1,t})}{\partial \log(P_{1,t})} \\
&= -\frac{1}{Q_{1,t}} \frac{\partial Q_{1,t}}{\partial \log(P_{1,t})} \\
&= -\frac{\gamma \Lambda_{1,\text{unspanned},t}}{M_{1,\text{unspanned},t}} \frac{\partial Q_{1,t}}{\partial \log(P_{1,t})} \\
&= \frac{\gamma \Lambda_{1,\text{unspanned},t}}{M_{1,\text{unspanned},t}} \left(\left(\frac{1}{\gamma \Lambda_{1,\text{unspanned},t}} \right) \left(-\frac{\partial M_{1,t}}{\partial \log(P_{1,t})} \right) \right. \\
&\left. + \left(\frac{1}{\gamma \Lambda_{1,\text{unspanned},t}} \right) \left(\frac{\partial B_{-1,t}}{\partial \log(P_{1,t}} \right) + \left(\frac{M_{1,\text{unspanned},t}}{\gamma \Lambda^2_{1,\text{unspanned},t}} \right) \left(\frac{\partial \Lambda_{1,\text{unspanned},t}}{\partial \log(P_{1,t})} \right) \right) \\
&= \left(\frac{1}{M_{1,\text{unspanned},t}} \right) \left(-\frac{\partial M_{1,t}}{\partial \log(P_{1,t})} \right) + \left(\frac{1}{M_{1,\text{unspanned},t}} \right) \left(\frac{\partial B_{-1,t}}{\partial \log(P_{1,t})} \right) \\
&+ \left(\frac{1}{\Lambda_{1,\text{unspanned},t}} \right) \left(\frac{\partial \Lambda_{1,\text{unspanned},t}}{\partial \log(P_{1,t})} \right) \\
&= \left(\frac{1}{M_{1,\text{unspanned},t}} \right) \left(-\frac{\partial M_{1,t}}{\partial \log(P_{1,t})} \right) + \left(\frac{1}{M_{1,\text{unspanned},t}} \right) \left(\frac{\partial B_{-1,t}}{\partial \log(P_{1,t}} \right) \\
&+ \left(\frac{1}{\Lambda_{1,\text{unspanned},t}} \right) \left(\frac{\partial \Lambda_{1,t}}{\partial \log(P_{1,t})} \right) - \left(\frac{1}{\Lambda_{1,\text{unspanned},t}} \right) \left(\frac{\partial \Lambda_{-1,t}}{\partial \log(P_{1,t})} \right).
\end{aligned}
$$

where

$$\Lambda_{-1,t} = B_{-1,t} \Lambda_{1,2,t}.$$

Thus, this calculation closely mirrors Proposition 2. The demand elasticity for asset 1 is broken into three components: the "main component", which reflects changes in the expected dollar return $M_{1,t}$, the "volatility component", related to changes in the variance $\Lambda_{1,t}$, and the "correlation component", which captures the changes in cross-asset exposure $B_{-1,t}$ and covariances. The final elasticity expression is:

$$\eta_{1,t} = \underbrace{\frac{1}{M_{1,\text{unspanned},t}}\left(-\frac{\partial M_{1,t}}{\partial \log(P_{1,t})}\right)}_{\text{main component}} + \underbrace{\frac{1}{\Lambda_{1,\text{unspanned},t}}\left(\frac{\partial \Lambda_{1,t}}{\partial \log(P_{1,t})}\right)}_{\text{volatility component}} + \underbrace{\frac{1}{M_{1,\text{unspanned},t}}\left(\frac{\partial B_{-1,t}}{\partial \log(P_{1,t})}\right) - \frac{1}{\Lambda_{1,\text{unspanned},t}}\left(\frac{\partial \Lambda_{-1,t}}{\partial \log(P_{1,t})}\right)}_{\text{correlation component}}.$$

Notice here that there is no "one plus" term that Equation (13) starts with. This is because we write demand here not in terms of portfolio weights, but as shares. Now we walk through this math in terms of returns, unspanned returns, and unspanned variances. Thus, we will now use excess returns in percentages, not dollars, as in the following:

$$r_{i,t+1} = \frac{P_{i,t+1} + D_{i,t+1} - R_{f,t}P_{i,t}}{P_{i,t}},$$

and thus

$$\begin{aligned} M_{i,t} &= P_{i,t}\mathbb{E}_t[r_{i,t+1}] = P_{i,t}\mu_{i,t} \\ \Lambda_{i,t} &= P_{i,t}^2\text{Var}_t[r_{i,t+1}] = P_{i,t}^2\sigma_{i,t}^2 \\ \Lambda_{i,j,t} &= P_{i,t}P_{j,t}\text{Cov}_t(r_{i,t+1}, r_{j,t+1}) = P_{i,t}P_{j,t}\sigma_{i,j,t} \end{aligned}$$

We can use the above to solve for these terms:

$$
\begin{aligned}
B_{-1,t} &= \frac{\Lambda_{1,2,t}}{\Lambda_{2,t}} = \frac{P_{1,t}P_{2,t}\sigma_{1,2,t}}{P_{2,t}^2\sigma_{2,t}^2} = \frac{P_{1,t}\sigma_{1,2,t}}{P_{2,t}\sigma_{2,t}^2} = \frac{P_{1,t}}{P_{2,t}}\beta_{-1,t} \\
B_{-2,t} &= \frac{\Lambda_{1,2,t}}{\Lambda_{1,t}} = \frac{P_{1,t}P_{2,t}\sigma_{1,2,t}}{P_{1,t}^2\sigma_{1,t}^2} = \frac{P_{2,t}\sigma_{1,2,t}}{P_{1,t}\sigma_{1,t}^2} = \frac{P_{2,t}}{P_{1,t}}\beta_{-2,t} \\
M_{1,t} &= P_{1,t}\mu_{1,t} \\
M_{2,t} &= P_{2,t}\mu_{2,t} \\
M_{1,\text{unspanned},t} &= M_{1,t} - B_{-1,t}M_{2,t} = P_{1,t}\mu_{1,t} - \frac{P_{1,t}\sigma_{1,2,t}}{P_{2,t}\sigma_{2,t}^2}P_{2,t}\mu_{2,t} = P_{1,t}\left(\mu_{1,t} - \frac{\sigma_{1,2,t}\mu_{2,t}}{\sigma_{2,t}^2}\right) \\
&= P_{1,t}\mu_{1,\text{unspanned},t} \\
M_{2,\text{unspanned},t} &= M_{2,t} - B_{-2,t}M_{1,t} = P_{2,t}\mu_{2,t} - \frac{P_{2,t}\sigma_{1,2,t}}{P_{1,t}\sigma_{1,t}^2}P_{1,t}\mu_{1,t} = P_{2,t}\left(\mu_{2,t} - \frac{\sigma_{1,2,t}\mu_{1,t}}{\sigma_{1,t}^2}\right) \\
&= P_{2,t}\mu_{2,\text{unspanned},t} \\
\Lambda_{1,\text{unspanned},t} &= \Lambda_{1,t} - B_{-1,t}\Lambda_{1,2,t} = P_{1,t}^2\sigma_{1,t}^2 - \frac{P_{1,t}\sigma_{1,2,t}}{P_{2,t}\sigma_{2,t}^2}P_{1,t}P_{2,t}\sigma_{1,2,t} \\
&= P_{1,t}^2\left(\sigma_{1,t}^2 - \frac{\sigma_{1,2,t}^2}{\sigma_{2,t}^2}\right) = P_{1,t}^2\sigma_{1,\text{unspanned},t} \\
\Lambda_{2,\text{unspanned},t} &= \Lambda_{2,t} - B_{-2,t}\Lambda_{1,2,t}^2 = P_{2,t}^2\sigma_{2,t}^2 - \frac{P_{2,t}\sigma_{1,2,t}}{P_{1,t}\sigma_{1,t}^2}P_{1,t}P_{2,t}\sigma_{1,2,t} \\
&= P_{2,t}^2\left(\sigma_{2,t}^2 - \frac{\sigma_{1,2,t}^2}{\sigma_{1,t}^2}\right) = P_{2,t}^2\sigma_{2,\text{unspanned},t}
\end{aligned}
$$

So we can use this to write:

$$
\begin{aligned}
Q_{1,t} &= \frac{M_{1,\text{unspanned},t}}{\gamma\Lambda_{1,\text{unspanned},t}} = \frac{P_{1,t}\mu_{1,\text{unspanned},t}}{\gamma P_{1,t}^2\sigma_{1,\text{unspanned},t}} = \frac{\mu_{1,\text{unspanned},t}}{\gamma P_{1,t}\sigma_{1,\text{unspanned},t}} \\
Q_{2,t} &= \frac{M_{2,\text{unspanned},t}}{\gamma\Lambda_{2,\text{unspanned},t}} = \frac{P_{2,t}\mu_{2,\text{unspanned},t}}{\gamma P_{2,t}^2\sigma_{2,\text{unspanned},t}} = \frac{\mu_{2,\text{unspanned},t}}{\gamma P_{2,t}\sigma_{2,\text{unspanned},t}}
\end{aligned}
$$

To calculate the elasticity of asset 1, we can first take the log of demand:

$$\log(Q_{1,t}) = \log\left(\frac{\mu_{1,\text{unspanned},t}}{\gamma\sigma_{1,\text{unspanned},t}}\right) - \log(P_{1,t})$$

Thus the elasticity is given by:

$$\begin{aligned}\eta_{i,t} &= -\frac{\partial \log(Q_{1,t})}{\partial \log(P_{1,t})} \\ &= 1 - \frac{\partial}{\partial \log(P_{1,t})} \log\left(\frac{\mu_{1,\text{unspanned},t}}{\gamma\sigma_{1,\text{unspanned},t}}\right)\end{aligned}$$

Thus we can walk through the same steps above to obtain the following:

$$\begin{aligned}\eta_{1,t} = \; & \underbrace{1 + \frac{1}{\mu_{1,\text{unspanned},t}}\left(-\frac{\partial \mu_{1,t}}{\partial \log(P_{1,t})}\right)}_{\text{main component}} \\ & + \underbrace{\left(\frac{1}{\sigma^2_{1,\text{unspanned},t}}\right)\frac{\partial \sigma^2_{1,t}}{\partial \log(P_{1,t})}}_{\text{volatility component}} + \underbrace{\frac{1}{\mu_{1,\text{unspanned},t}}\mu_{-1,t}\frac{\partial \beta_{-1,t}}{\partial \log(P_{1,t})} - \left(\frac{1}{\sigma^2_{1,\text{unspanned},t}}\right)\frac{\partial \sigma^2_{-1,t}}{\partial \log(P_{1,t})}}_{\text{correlation component}}.\end{aligned}$$

This matches Proposition 2.