The Lost Capital Asset Pricing Model

Daniel Andrei∗ Julien Cujean† Mungo Wilson‡

June 14, 2018

Abstract

A flat Securities Market Line is not evidence against the CAPM. In a rational-expectations economy in which markets are inefficient, the CAPM holds but is rejected empirically (Type I Error). There exists an information gap between the empiricist and the average investor who clears the market. The CAPM holds unconditionally for the investor, but appears flat to the empiricist who uses the correct unconditional market proxy. This distortion is empirically substantial and offers a new interpretation of why “Betting Against Beta” works: BAB really bets on true beta. The empiricist retrieves a stronger CAPM on macroeconomic announcement days.

∗UCLA Anderson Graduate School of Management, 110 Westwood Plaza, Suite C420, Los Angeles, CA 90095, USA; +1 (310) 825-3544; daniel.andrei@anderson.ucla.edu; www.danielandrei.info.
†CEPR and Robert H. Smith School of Business, University of Maryland, 4466 Van Munching Hall, College Park, MD 20742, USA; +1 (301) 405-7707; jcujean@rhsmith.umd.edu; juliencujean.com.
‡Said Business School and Oxford-Man Institute, Oxford University, United Kingdom; +44 (0)1865 288914; mungo.wilson@sbs.ox.ac.uk.
1 Introduction

There is a growing tension between the theory of financial economics and its application. The Capital Asset Pricing Model, a theoretical pillar of modern finance, fails in empirical tests.\(^1\) The consensus among economists is that beta does not explain expected returns, largely shaping the view that the CAPM does not hold. But a flagrant affront to this view is that the CAPM remains to this day the model that investors and firms most widely use.\(^2\) Adding to the controversy, the CAPM does hold on particular occasions, e.g., on announcement days, or at night.\(^3\) Why do economists keep rejecting a theory that practitioners are not willing to abandon?

This paper explores the idea that the CAPM is rejected by mistake. We argue that an empiricist may incorrectly reject the CAPM when in fact it is the correct asset pricing model—the empiricist commits a Type I error. There are many reasons not to believe the CAPM is the correct canonical asset pricing model; and there are as many ways it could fail empirically. In this paper, we present a situation in which the CAPM holds, but it fails empirically in one specific way: the empiricist perceives a “flat” Securities Market Line, which becomes steeper occasionally, e.g., when public information is released.

The mechanism we propose is relevant when investors have information that the empiricist does not (the Hansen and Richard (1987) critique\(^4\)) and that market prices do not fully reflect. Because investors must hold the portfolio of aggregate wealth in equilibrium, the composition of this portfolio is a central determinant of equilibrium prices, in addition to investors’ information. Allocation of risk in markets creates noise in the composition of this portfolio (Black, 1986; Grossman, 1995), which impairs informational efficiency. In this economy, noise makes the information gap between investors and the empiricist matter for the CAPM test, even when the empiricist uses the correct market proxy. We construct an economic representation of this information gap and determine how the empiricist perceives the SML.

Suppose an empiricist (an outside observer) observes the time series of realized excess returns for a large number of assets. Using the law of total covariance, the unconditional variance-covariance matrix of excess returns the empiricist computes based on historical data,\(^1\)

---

\(^1\)See Fama and French (2004) for a comprehensive review.

\(^2\)The CAPM remains the model most widely used to make investment decisions (Berk and Van Binsbergen, 2016; Barber, Huang, and Odean, 2016) and to compute the cost of capital (Graham and Harvey, 2001).

\(^3\)Savor and Wilson (2014) document a strong relationship between expected returns and betas on days when news about inflation, unemployment, or FOMC interest rate decisions is scheduled to be announced. Hendershot, Livdan, and Rösch (2018) document a strong relationship when the market is closed (at night). Ben-Rephael, Carlin, Da, and Israelsen (2017) document that the CAPM performs better when institutions demand more information.

\(^4\)See Roll (1978) and Dybvig and Ross (1985) for a similar argument.
\[ \mathbb{V}[R], \text{ can be decomposed into the observed variation, } \mathbb{V}[\mathbb{E}[R|\mathcal{F}]], \text{ “explained” by the information } \mathcal{F} \text{ of the average investor, and a remaining “unexplained” component, } \mathbb{E}[\mathbb{V}[R|\mathcal{F}]]: \]

\[ \mathbb{V}[R] = \mathbb{V}[\mathbb{E}[R|\mathcal{F}]] + \mathbb{E}[\mathbb{V}[R|\mathcal{F}]]. \]  

This decomposition formalizes the notion of *informational distance* between the empiricist and the average investor.

The true betas in the economy—investor’s betas—are solely based on the unexplained component of the variance. Empiricist’s betas, instead, are based on the total variance, \[ \mathbb{V}[R]. \]

True and measured betas thus differ in proportion to the informational distance in Eq. (1). We emphasize that an unconditional CAPM relation always holds in this paper. The classic argument that variation in beta invalidates an unconditional CAPM relation (Jagannathan and Wang, 1996; Lewellen and Nagel, 2006) does not apply in our context. Nor is the CAPM correctly rejected when it does not hold (Merton, 1987). In this paper, the CAPM is the correct model but the empiricist mis-estimates unconditional betas.

We build our argument in a rational-expectations model of informed trading in which a continuum of mean-variance investors trade multiple assets based on private and public information. A necessity for private information to be relevant in this framework is that the market portfolio contain noise (Grossman and Stiglitz, 1980). We show that the Hansen and Richard (1987) critique only matters in our model when the market portfolio is noisy. Noise leads the informational distance in Eq. (1) to impair the empiricist’s ability to recover the unconditional CAPM relation.

Using a proxy for the market portfolio is a necessity for carrying out a test of the CAPM. That CAPM tests fail when the market proxy is not mean-variance efficient is certainly true (Roll and Ross, 1994) and is not our question. Our question is how does the empiricist reject the CAPM using a market proxy that is theoretically correct. We thus let the empiricist choose a proxy that is unconditionally mean-variance efficient from the perspective of the “average investor,” a fictitious agent who holds the market portfolio and whose beliefs define the market consensus. The empiricist, who has a coarser information set and thus faces more uncertainty than the average investor, perceives the comovement among assets differently and rejects the CAPM. Hence, although empiricist’s proxy is mean-variance efficient for the average investor, it is not so for the empiricist.

The main result is that in equilibrium the empiricist perceives a “flat” CAPM relation. The average investor sees the true SML—a line that crosses the origin with the market risk premium as its slope—but the empiricist’s SML is flatter. Figure 1 illustrates this distortion, plotting the true SML (the dashed line) and the perceived SML (the solid line). The
Figure 1: CAPM distortion. This figure illustrates the main result of the paper. The perceived SML is flatter than the actual SML in equilibrium. The black dashed line and the red solid line show the true and perceived SML. \( \mathbf{M} \) represents the market portfolio.

Informational distance in Eq. (1) amplifies the dispersion in empiricist’s betas \( (\tilde{\beta}_1 \text{ and } \tilde{\beta}_2) \) relative to true betas \( (\beta_1 \text{ and } \beta_2) \). However, all betas (correct or incorrect) must average to one. The market beta thus becomes the “center of gravity” around which empiricist’s betas inflate away from a value of one. Assets with a beta higher than one appear riskier than they really are, whereas assets with a beta lower than one appear safer than they really are. Since the empiricist and the investor agree on what unconditional returns are, empiricist’s SML rotates clockwise around the market portfolio (denoted by \( \mathbf{M} \)), which flattens its slope and creates a positive intercept.

Interestingly, this result—true betas are shrunk towards one relative to measured betas—corresponds to the way practitioners adjust beta estimates (e.g., “ADJ BETA” on Bloomberg terminals). This adjustment in our model and in practice has different origins. Practitioners use it to reduce sampling biases (Vasicek, 1973), which are absent in our model. However, we can still compare how much shrinkage is reasonable in our model with how much shrinkage finance textbooks recommend (e.g., Berk and DeMarzo, 2007). This requires time series of expected returns, which are typically difficult to measure because they are determined by information that investors possess but that empiricists do not. Fortunately, option prices tell us what the market thinks about future returns, thus providing a proxy for what the average investor’s expectations are (e.g, Martin, 2017). Back-of-the-envelope calculations suggest that the distortion can be large, and that shrinkage in practice is too conservative.

That this distortion has plausibly large empirical magnitudes offers an alternative interpretation of existing findings in the literature. Under our theory, betting against measured
beta (Frazzini and Pedersen, 2014) is really betting on true beta. Furthermore, in the eyes of
the empiricist, the distortion in beta estimates creates the illusion of an idiosyncratic volatility puzzle—stocks with high idiosyncratic risk have implausibly low returns (Ang, Hodrick, Xing, and Zhang, 2006, 2009).

Eliminating this distortion is empirically daunting. There will always be asset-pricing relevant information that the empiricist ignores. Individual investors possess private information that prices do not fully reveal. And the empiricist will never be able to state confidently that she has controlled for all publicly-available information. Nevertheless, opportunities exist for the empiricist to reduce the distortion in betas and retrieve a stronger CAPM. One such opportunity is to estimate the CAPM across different market cycles, e.g., trading and non-trading hours. The amount of informational trading and of mispricing associated with noise is likely distinct across cycles (French and Roll, 1986). Our theory predicts a stronger CAPM when there is less noise and less informational trading, i.e., at night (Hendershott et al., 2018).

Public announcements are another opportunity for the empiricist to reduce the distortion in betas. Announcements provide macroeconomic information about payoffs, which may reduce the informational distance between investors and the empiricist. They also involve macroeconomic uncertainty that may raise the risk premium prior to their release. To explore this possibility, we allow the empiricist in our model to estimate the CAPM relation conditioning on days when investors observe public announcements (Savor and Wilson, 2014). We show that the informational improvement and the increase in the risk premium jointly lead to a distinctly steeper CAPM relation on announcement days.

There are several established explanations for the finding that the SML is too flat, some of which go back to the 1970’s. None directly result from the informational distance in Eq. (1). We believe that the fleeting appearance of the CAPM on announcement days—let alone its pervasive application in practice—licenses a new look at the finding that the SML is too flat. Assuming the CAPM holds unconditionally, we argue that an empiricist may incorrectly reject it using the correct market proxy. This is different from classic arguments whereby the CAPM either does not hold or only holds conditionally (Dybvig and Ross, 1985; Hansen and Richard, 1987; Merton, 1987; Jagannathan and Wang, 1996), or arguments whereby the market proxy is incorrect (Roll, 1977; Stambaugh, 1982; Roll and Ross, 1994).

Section 2 presents our main result in a static model and provides intuition about the

---

5These explanations include leverage constraints (Black, 1972; Frazzini and Pedersen, 2014), inflation (Cohen, Polk, and Vuolteenaho, 2005), disagreement (Hong and Sraer, 2016), preference for volatile, skewed returns (Kumar, 2009; Bali, Cakici, and Whitelaw, 2011), market sentiment (Antoniou, Doukas, and Subrahmanyam, 2015), stochastic volatility (Campbell, Giglio, Polk, and Turley, 2012), and benchmarking of institutional investors (Baker, Bradley, and Wurgler, 2011; Buffa, Vayanos, and Woolley, 2014).
distortion in beta estimates, along with back-of-the-envelope calculations of the distortion. Section 3 reinterprets anomalies in light of this result and examines the distortion of the SML in a dynamic model with a periodic public announcement. Section 4 concludes.

2 The Empiricist’s Perception of the SML

We propose a new answer to an old asset-pricing question: why is the Securities Market Line flat? Suppose a true CAPM relation holds under a representative investor’s information set. The empiricist, who does not observe this information set (Hansen and Richard, 1987), rejects the CAPM when in fact it is the correct asset-pricing model. We first clarify under which conditions the informational distance between the empiricist and the investor matters. We then explore its implications for CAPM tests in an equilibrium model with dispersed information. Our main result is that the empiricist views low-beta assets as less risky than they actually are and high-beta assets as riskier than they actually are—the empiricist’s SML looks flat. We argue that this distortion is empirically substantial.

2.1 Type I Error: Rejecting the CAPM when it Holds True

To illustrate how our mechanism operates, we go through a typical exercise of deriving the CAPM. Consider a one-period economy populated by a representative investor who derives monotone increasing utility, \( U(W) \), from consuming her wealth \( W \) in period one. For simplicity, suppose the investor has zero initial wealth (what she invests in the \( N \) stocks available in supply \( M \) and paying excess returns \( R \) she must borrow from a risk-free bond with gross return normalized to 1). In equilibrium, the supply of stocks must equal investor’s demand and thus \( M \) defines the market portfolio.

Investor’s first-order condition for optimal portfolio choice leads to the standard asset-pricing equation:

\[
\mathbb{E}[U(W)(M'R)|\mathcal{F}] = 0,
\]

where \( \mathcal{F} \equiv \{M, X\} \) denotes the investor’s information set; it contains all information that is relevant for asset pricing. This information must include the market portfolio \( M \in \mathcal{F} \), as well as additional asset-pricing relevant information, which we denote by \( X \).

For the CAPM to be the correct asset-pricing model we need appropriate modeling assumptions. Among possible frameworks we choose a CARA-normal structure. We assume that asset-pricing information \( X \) and the vectors of market weights \( M \) and excess returns \( R \) are jointly Gaussian. Under this assumption, Stein’s lemma applies and conditional covari-
ances are nonrandom. We also restrict preferences, $U(\cdot)$, to be exponential so that the market price of variance risk is a known constant, $\gamma$, the investor’s coefficient of absolute risk aversion. This structure is also analytically convenient for incorporating disperse information, which we will do shortly.

The CAPM relation is then derived in a few standard steps (see Appendix A.1). First, rearrange Eq. (2) using covariance decomposition and apply Stein’s lemma:

$$\mathbb{E}[R|\mathcal{F}] = \gamma \mathbb{V}[R|\mathcal{F}]M. \tag{3}$$

This relation must also hold for the market portfolio:

$$\mathbb{E}[R|\mathcal{F}] = \beta_M \mathbb{E}[R_M|\mathcal{F}], \tag{4}$$

where $R_M \equiv M'R$ denotes excess returns on the market portfolio and $\beta_M$ is the vector of conditional betas. Eq. (4) describes the conditional CAPM relation. Taking unconditional expectations and using covariance decomposition yields:

$$\mathbb{E}[R] = \beta \mathbb{E}[R_M] + \text{Cov}[\beta_M, \mathbb{E}[R_M|\mathcal{F}]] + (\mathbb{E}[\beta_M] \mathbb{E}[R_M] - \beta \mathbb{E}[R_M]), \tag{5}$$

where $\beta$ is the vector of unconditional betas, $\bar{M} = \mathbb{E}[M]$, and $R_M' = \bar{M}'R$.

Even if the CAPM holds conditionally, it need not hold unconditionally. Beyond the market, additional “factors” emerge through variation in conditional betas (Jagannathan and Wang, 1996). These factors capture the extent to which the conditional and unconditional CAPM differ. Unless these factors are together zero, the unconditional CAPM fails. Lewellen and Nagel (2006) argue that the deviation caused by these factors is empirically small. In the present model, this deviation is zero. Specifically, the risk-return tradeoff in Eq. (3) implies that the additional factors cancel each other out, and arguments made in Jagannathan and Wang (1996) do not apply in our setup:

$$\mathbb{E}[R] = \beta \mathbb{E}[R_M] + \text{Cov}[\beta_M, \mathbb{E}[R_M|\mathcal{F}]] - \gamma \text{Cov}[\beta_M, \mathbb{V}[R_M|\mathcal{F}]] \approx \beta \mathbb{E}[R_M]. \tag{6}$$

We conclude that the unconditional CAPM holds in our setup. The question we now ask is whether the empiricist reaches the same conclusion. Note that this is different from testing the CAPM in a model where it does not hold (e.g., Merton, 1987); rejecting the CAPM when it is incorrect would be the correct conclusion. Instead, we argue that the empiricist incorrectly rejects the CAPM when it holds—she commits a type I error.
While the empiricist does not observe the information set $\mathcal{F}$ of the representative investor, the law of iterated expectations certainly allows her to condition Eq. (2) down to an unconditional expectation:

$$E[U_W(M'R)R] = 0.$$  \hspace{1cm} (7)

That, however, is not much good unless the empiricist knows the pricing kernel $U_W(M'R)$. The pricing kernel depends functionally on the market portfolio $M$. The market portfolio contains noise, which the empiricist does not observe. Nevertheless, for the purposes of our argument, we assume that the empiricist correctly measures the unconditional market portfolio $\bar{M}$, on which the unconditional CAPM is based (Eq. 6). Although this portfolio is the correct proxy for the test, the empiricist concludes that $\bar{M}$ is not mean-variance optimal and incorrectly rejects the CAPM:

$$E[U_W(\bar{M}'R)R] = E[U_W(M'R)](\mathbb{V}[R] - \mathbb{V}[R|\mathcal{F}])\bar{M} \neq 0.$$  \hspace{1cm} (8)

Noise in the market portfolio makes the informational distance between the investor and the empiricist matter—even when the empiricist uses the correct market proxy for the test. (This distance matters even when the empiricist measures returns at higher frequency—see Appendix A.2). As Eq. (8) shows, the empiricist’s proxy violates optimality in proportion to this distance. However, although this equation shows that the informational distance distorts the test, it does not say how. To place economic restrictions on the resulting distortion of the CAPM test, we need a representation of the distance $\mathbb{V}[R] - \mathbb{V}[R|\mathcal{F}]$ in an equilibrium model of returns. This is the goal of the next section.

2.2 An Equilibrium Model of Excess Returns

We now build a model of how investors form expectations, imposing an equilibrium structure on excess returns. Consider the one-period economy of Section 2.1 in which the market consists of one risk-free asset with gross return normalized to 1 and $N$ risky stocks indexed by $n = 1, ..., N$. Suppose the risky stocks have payoffs $D$ realized at the liquidation date (time 1). These payoffs are unobservable at the trading date (time 0) and have a common
factor structure:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_N \end{bmatrix} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} F + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix} \equiv \Phi F + \epsilon. \quad (9)$$

The common factor $F$ and each stock-specific component $\epsilon_j$ are independently normally distributed with means zero and precisions $\tau_F$ and $\tau_\epsilon$. Without loss of generality, we assume that the cross-sectional average of the loadings of assets’ payoffs on the common factor is positive, i.e., $\Phi \equiv N^{-1} \sum_{n=1}^{N} \phi_n > 0$.

The economy is populated with a continuum of investors indexed by $i \in [0, 1]$, who choose their portfolio at time 0 and derive utility from terminal wealth with constant absolute risk aversion coefficient $\gamma$. Investors know the structure of realized payoffs in Eq. (9), but do not observe the common factor. Each investor $i$ forms expectations about $F$ based on both a private signal $V_i$ and a public signal $G$:

$$V^i = F + v^i \quad (10)$$
$$G = F + v. \quad (11)$$

Signal noises $v$ and $v^i \perp v$, $\forall i$ are unbiased and independently normally distributed with precisions $\tau_G$ and $\tau_v$, respectively.

In this economy informational efficiency fails—equilibrium prices do not fully reveal investors’ private information about the common factor $F$. Prices change to reflect new information about final payoffs, but they also change for reasons unrelated to information, e.g., endowments shocks, preference shocks, or private investment opportunities. To model uninformative price changes, we assume that the supply of stocks, $M \equiv [M_1 \ldots M_N]'$, is noisy (Grossman and Stiglitz, 1980). This assumption makes the market portfolio unobservable both to investors and to the empiricist. On average, each stock has equal weight $1/N$ in the market portfolio and each weight is normally and independently distributed across stocks with precision $\tau_M$.

This economy relies on several simplifying assumptions. We have assumed that payoffs in Eq. (9) are driven by a single factor, as opposed to multiple factors. Stocks only differ

---

6There are many ways to endogenize uninformative trading (noise). For instance, Wang (1994) analyzes a model in which investors have private investment opportunities. Alternatively, investor specific endowment shocks can also generate uninformed trading. Another possibility is to introduce a mass of hedgers endowed with income risk that is correlated with assets’ payoffs (Farboodi and Veldkamp, 2017). These alternatives would unnecessarily complicate the analysis, without bringing additional economic insights.
according to their loading $\Phi$ on this common factor—they have equal weight (or size) in the market portfolio on average; their sizes and their idiosyncratic noises have equal precision. Similarly, because the model is static, public and private signals have identical timing. These simplifications serve our purpose of isolating the main result.

We solve for a linear equilibrium of the economy in which prices satisfy

$$ P = \alpha F + g G + \xi M, \quad (12) $$

where $\alpha$ and $g$ are $N$–dimensional vectors and $\xi$ is a $N \times N$ matrix, all of which are determined in equilibrium by imposing market clearing. Because in this framework returns are not normally distributed, a convention in the literature is to work with price changes instead (e.g., Dybvig and Ross, 1985). We follow this convention and refer to price changes, $R \equiv D - P$, as “excess returns.”

Each investor $i$ forms expectations about excess returns based on her information set:

$$ F^i = \{V^i, G, P\}. \quad (13) $$

Because private signals $V_i$ all have identical precision, and the signal $G$ and prices $P$ are public, each investor $i$ forecasts the common factor $F$ with the same precision:

$$ \tau \equiv \nabla[F \mid F^i]^{-1} = \tau_F + \tau_v + \tau_G + \tau_P \Phi' \Phi. \quad (14) $$

The last term in Eq. (14) is the sum of squared signal-to-noise ratios over all prices, and $\tau_P$ is a scalar that measures price informativeness. Prices have explicit solutions that we provide in the proposition below (its proof is given in Appendix A.3).

**Proposition 1. (Equilibrium)** There exists a unique linear equilibrium in which prices take the linear form in Eq. (12) and are explicitly given by

$$ P = \frac{\tau - \tau_F - \tau_G}{\tau} \Phi F + \frac{\tau_G}{\tau} \Phi G - \left( \frac{\gamma + \sqrt{\tau_M \tau_P}}{\tau} \Phi \Phi' + \frac{\gamma}{\tau_e} I_N \right) M, \quad (15) $$

where $I_N$ is the identity matrix of dimension $N$. The precision $\tau$ is defined in (14) and the scalar $\tau_P$ is the unique positive root of the cubic equation:

$$ \tau_P \left[ \tau_F + \tau_v + \tau_G + (\tau_P + \tau_e) \Phi' \Phi \right]^2 \gamma^2 = \tau_M \tau_e^2 \tau_v^2. \quad (16) $$

In this economy investors have different perceptions of the mean-variance frontier, because each investor $i$ observes a different information set $F^i$. Hence, not only it is impossible for
them to observe and hold the market portfolio \( M \), but from their perspective it would not even be mean-variance optimal to do so. Holding the market portfolio is both mean-variance efficient and feasible only for the average investor, a fictitious investor who defines the consensus (average) beliefs and clears the market. We denote the beliefs of this average investor by \(( \hat{E}[\cdot], \hat{V}[\cdot] )\) and describe them below.

Since the precision on the common factor in Eq. (14) is identical for all investors, they hold the same posterior variance of excess returns. Thus, the posterior variance of excess returns from the perspective of the average investor is:

\[
\hat{V}[R] \equiv \forall[R \mid \mathcal{F}^i] = \frac{1}{\tau} \Phi \Phi' + \frac{1}{\tau \epsilon} \mathbf{1}_N, \quad \forall i \in [0, 1].
\]  

(17)
The consensus beliefs are further defined by averaging over investors’ expectations:

\[
\hat{E}[R] \equiv \int_i E[R \mid \mathcal{F}^i] di.
\]  

(18)

Because the average investor clears the market, the market-clearing condition relates endogenously the consensus beliefs and posterior variance of excess returns:

\[
\hat{E}[R] = \gamma \hat{V}[R] M,
\]  

(19)

which leads to the solution for equilibrium prices in Proposition 1. In other words, since the average investor clears the market, she must find it mean-variance optimal to hold the market portfolio \( M \). Furthermore, Eq. (19) defines the true capital market line, which represents the particular form Eq. (3) takes in this model. By the law of iterated expectations, it follows as a corollary that a CAPM relation holds unconditionally for the average investor.

**Corollary 1.1. (CAPM)** In this model, an unconditional CAPM relation holds:

\[
E[R] = \frac{\text{Cov}[R, R_M]}{\hat{V}[R_M]} E[R_M] = \beta E[R_M],
\]  

(20)

where \( \beta \) is a \( N \)-dimensional vector of betas. Denoting by \( \mathbf{1} \) a \( N \)-dimensional vector of ones, \( \text{Cov}[R, R_M] \) is the vector of covariances between excess returns of individual stocks and excess returns on the market, \( E[R_M] \) is the unconditional expected excess return on the market, and \( \hat{V}[R_M] \) is the variance of excess returns on the market:

\[
\hat{\text{Cov}}[R, R_M] = \frac{1}{N} \hat{V}[R] \mathbf{1}, \quad E[R_M] = \frac{1}{N} \mathbf{1}' E[R], \quad \hat{V}[R_M] = \frac{1}{N^2} \mathbf{1}' \hat{V}[R] \mathbf{1}.
\]  

(21)
Equation (20) is the unconditional CAPM. In this economy, computing actual betas, $\beta$, only requires knowing the variance of excess returns measured by the average investor, $\hat{\sigma}(R)$.

### 2.3 The Empiricist’s View in Equilibrium

Given the equilibrium representation of the CAPM in Corollary 1.1, we now investigate how the informational distance in Eq. (8) affects the empiricist’s own representation of the CAPM. For simplicity, we assume the empiricist can only compute unconditional moments from realized excess returns.\(^7\)

The starting point is the law of total covariance in Eq. (1), which allows us to formalize the informational distance between the average investor and the empiricist:

$$
\sigma(R) - \hat{\sigma}(R) = \sigma[E[R,F^i)] = \sigma[\hat{E}[R] + \Phi \frac{\tau v}{\tau^2} v_i] = \sigma[\hat{E}[R]] + \frac{\tau v}{\tau^2} \Phi \tau.
$$

(22)

The first equality states that the empiricist perceives additional variation in realized returns relative to investors of the model—she observes none of the information in $F^i$. Because this information is heterogeneous across investors, their expectations differ from consensus beliefs by the idiosyncratic noise in their private signal (the meaning of the second equality); accordingly, the last equality decomposes the additional variation the empiricist perceives into variation in consensus beliefs and variation in investors’ dispersed information.

The empiricist perceives additional variation relative to an investor $i$, because from investor’s perspective returns are predictable. Were returns unpredictable, all unconditional variation in excess returns, $\sigma(R)$, would be due to conditional variation $\hat{\sigma}(R)$, and none to variation in expected returns. But in this economy noise impairs informational efficiency and returns are predictable. As a result, from the perspective of the empiricist, conditioning down distorts the measurement of unconditional variation in realized excess returns and thus in unconditional betas.

Without adding economic content to Eq. (22), however, this statistical relation does not say how it distorts empiricist’s betas. To augment this statistical decomposition with an economic argument, we use the equilibrium relation (19). This provides an endogenous link between the variation in excess returns measured by the empiricist, $\sigma(R)$, and the variation in excess returns measured by the average investor, $\hat{\sigma}(R)$. We characterize this link in the following Lemma, and provide its proof in Appendix A.4.

---

\(^7\)We relax this assumption in Appendix A.6, allowing the empiricist to control for all publicly available information—i.e., prices and the public signal $G$—when testing the CAPM. The results hold even after controlling for all publicly available information (the empiricist cannot control for private information).
Lemma 1. In equilibrium, the informational distance between the average investor and the empiricist satisfies:

$$V[R] - \hat{V}[R] = \frac{\gamma^2}{\tau_M \tau_e} \hat{V}[R] + \frac{\kappa}{\tau} \Phi \Phi',$$

where $\kappa$ is a strictly positive scalar:

$$\kappa \equiv \frac{\gamma^2}{\tau_M} \left( \frac{1}{\tau_e} + \frac{\Phi' \Phi}{\tau} \right) + \frac{\tau_v}{\tau} > 0. \quad (24)$$

Absent noise in the market portfolio ($\tau_M \to \infty$), the informational distance in Eq. (23) vanishes. In equilibrium, variation in consensus beliefs arises because they move with the market portfolio $M$ (see Eq. 19). Eliminating noise removes variation in consensus beliefs. Similarly, eliminating noise allows investors to gain perfect knowledge of the common factor (i.e., $\tau \to \infty$), which makes private information irrelevant. Noise (informational inefficiency) thus creates the informational gap in Eq. (23).

Remarkably, the informational distance in Eq. (23) is determined in equilibrium by a unique, positive coefficient $\kappa$. The first term implies that the empiricist’s covariance is an inflated version of investors’ covariance. The second term distorts the variances of, and the covariances between, individual stocks. Further multiply Eq. (23) with $M$ to obtain:

$$\text{Cov}[R, R_M] = \left( 1 + \frac{\gamma^2}{\tau_M \tau_e} \right) \text{Cov}[R, \overline{R_M}] + \frac{\Phi \kappa}{\tau} \Phi. \quad (25)$$

All covariances between excess returns of individual stocks and excess returns on the market are inflated through the first term on the right hand side. This amplification effect strengthens with investors’ risk aversion, with the amount of noise in the market portfolio, and with the amount of noise in stock-specific payoffs. The second term on the right hand side shows that covariances with the market are inflated comparatively more for stocks with a stronger loading on the common factor. It follows that the empiricist observes higher betas for these stocks. Since betas must average to one, it also follows that the empiricist observes lower betas for the stocks with a weaker loading on the common factor.

The vector of betas that the empiricist estimates from realized returns is:

$$\hat{\beta} \equiv \frac{\text{Cov}[R, R_M]}{V[R_M]}. \quad (26)$$

As in Section 2.1, we allow the empiricist to use the correct proxy for the market portfolio, its average $\overline{M}$. This proxy is the market portfolio that the average investor finds mean-variance
optimal to hold unconditionally (Corollary 1.1). We then use Lemma 1 to determine how the empiricist perceives the CAPM relation. Theorem 1 is the main result of the paper (its proof is given in Appendix A.5).

**Theorem 1. (CAPM tests based on realized returns)** In the eyes of the empiricist, the expected excess return on each asset \( n \in \{1, 2, \ldots, N\} \) and on the market satisfy the relation:

\[
E[R_n] = \frac{\delta}{1 + \delta} (1 - \tilde{\beta}_n) E[R_M] + \beta_n E[R_M].
\]

(27)

In equilibrium the empiricist’s vector of betas, \( \tilde{\beta} \), and the average investor’s vector of betas, \( \beta \), both net of their average (their average is one) satisfy the proportionality relation:

\[
\tilde{\beta} - \mathbf{1} = (1 + \delta)(\beta - \mathbf{1}),
\]

(28)

where the strictly positive coefficient \( \delta \) measures the magnitude of the distortion in Eq. (27):

\[
\delta \equiv \frac{\kappa/N}{\tau_e \mathbb{V}[R_M]} = \frac{1}{N \mathbb{V}[R_M]} \left[ \gamma^2 \tau_e \left( \frac{1}{\tau_e} + \frac{\Phi' \Phi}{\tau} \right) + \tau_v \frac{\tau_v}{\tau} \right] > 0.
\]

(29)

The empiricist perceives mispricing (a non-zero alpha) for all stocks, except those that have a beta of one. Eq. (27) shows that low-beta assets (\( \tilde{\beta}_n < 1 \)) earn a positive unconditional alpha, whereas high-beta assets (\( \tilde{\beta}_n > 1 \)) earn a negative unconditional alpha. To see how perceived mispricing distorts the empiricist’s view of the SML, write (27) as:

\[
E[R_n] = \frac{\delta}{1 + \delta} E[R_M] + \tilde{\beta}_n \frac{E[R_M]}{1 + \delta},
\]

(30)

which resembles the zero-beta CAPM (Black, 1972). The first term implies that the empiricist perceives a SML with a positive intercept. The second term states that the perceived SML is flatter than the true SML of Corollary 1.1. Equations (27) and (30) describe the biggest failure of the CAPM (e.g., Black, Jensen, and Scholes, 1972, and the literature that followed): the high returns enjoyed by many apparently low-beta assets and the high intercept of the SML.

Figure 1 illustrates the SML distortion implied by Eq. (30). The empiricist’s SML rotates clockwise around the market portfolio, which flattens its slope and creates a positive intercept. How much flatter the SML is, and how large its intercept is, depends on the magnitude of the coefficient \( \delta \) in Eq. (29). This coefficient is proportional to \( \kappa \), which determines the size of the informational distance in Eq. (23). Hence, the larger this informational
distance is, the flatter empiricist’s SML. Likewise, absent noise in the market portfolio, the informational distance vanishes and so does the distortion in the SML.

The main result is that, in equilibrium, true betas are shrunk towards one relative to empiricist’s betas. The “degree of shrinkage” is determined by a unique coefficient, $\delta$, which adjusts the empiricist’s betas towards true betas. Interestingly, Eq. (28) is identical to the Bayesian estimator proposed by Vasicek (1973), an estimator that is popular in the financial industry (“ADJ BETA” on Bloomberg terminals). We emphasize, however, that the result of Theorem 1 is not due to sampling error. Nor is this result a standard attenuation bias, which commonly plagues the second pass cross-sectional regression in the Fama and MacBeth (1973) method. Rather, in equilibrium shrinkage in betas arises as a joint consequence of informational inefficiency and the Hansen and Richard (1987) critique.

The magnitude of the SML distortion in Eq. (29) depends not only on degree of informational inefficiency (measured by $1/\tau_M$), but also on investors’ risk aversion $\gamma$, on the degree of heterogeneity in assets’ loadings on the common factor (measured by $\Phi^T\Phi$), on the amount on noise in stock-specific payoffs (measured by $1/\tau_\epsilon$), and on the precision of private information $\tau_v$. These parameters, together or separately, can increase the magnitude of the distortion. The question that remains is, what is the empirical magnitude of the SML distortion? In other words, what is an empirically plausible value for the coefficient $\delta$ in Theorem 1? This matter is the subject of the next section.

### 2.4 Empirical magnitude of the distortion in beta

This section provides a back-of-the-envelope calculation of the distortion in beta estimates. Consider the unconditional beta of any individual security $n$, as computed by the empiricist from realized excess returns on the asset and on the market. We decompose this beta using the law of total covariance in Eq. (22):

$$\tilde{\beta}_n = \frac{\text{Cov}[R_n, R_M]}{\sqrt{\text{Var}[R_M]}} \approx \frac{\text{Cov}[R_n, R_M]}{\sqrt{\text{Var}[R_M]}} + \frac{\text{Cov}[\tilde{E}[R_n], \tilde{E}[R_M]]}{\sqrt{\text{Var}[R_M]}}. \tag{31}$$

The law of total covariance further implies that variation in investors’ private information causes the empiricist to perceive additional variation in realized returns (the last term in

---

8This linear adjustment has been first proposed by Blume (1971) (due to mean reversion of betas over time) and then by Vasicek (1973) (due to measurement error). See Bodie, Kane, and Marcus (2007), Berk and DeMarzo (2007) among others. Levi and Welch (2017) give best-practice advice for beta-shrinkage.

9In Vasicek (1973), the degree of adjustment depends on the sample size and converges to zero as the sample size increases. Similarly, Shanken (1992) shows that the attenuation bias becomes negligible as the length of the sample period grows indefinitely (see also Jagannathan and Wang, 1998; Shanken and Zhou, 2007; Kan, Robotti, and Shanken, 2013). In our case, the adjustment is necessary even in infinite samples.
Eq. 22). However, empirically disentangling this source of variation from variation in consensus beliefs requires observing data on individual investors’ information and this data is unavailable. Since our purpose is simply to approximate the empirical magnitudes of the distortion in betas, we abstract from this source of variation in realized returns. As a result, the second equality above holds as an approximation. We also assume that the covariance \( \bar{\text{Cov}}[R_n, R_M] \) is nonrandom (i.e., that betas of securities are not moving over time).\(^{10}\)

We can further decompose the approximation above as follows:

\[
\beta_n \approx \frac{\text{Cov}[R_n, R_M]}{\text{Var}[R_M]} \approx 1 - R^2
\]

\[
\beta_{E,n} = \frac{\text{Cov}[\bar{R}_n, \bar{R}_M]}{\text{Var}[\bar{R}_M]} \approx \frac{\text{Var}[\bar{R}_M]}{\text{Var}[R_M]} \approx R^2.
\]

The empiricist’s beta is a weighted average of two terms. The first term is the true beta of security \( n \) as measured by the average investor. The second term, \( \beta_{E,n} \), has a “beta-like” structure; it is a term that we could compute by means of simple (OLS) regressions if only we observed time series of expected excess returns on stock \( n \) and the market. Both terms are weighted by \( R^2 \), which represents the coefficient of determination from regressing excess returns of the market portfolio on the information set of the average investor.

Although the relation in Eq. (32) is merely a statistical decomposition, when combined with our theory it becomes a testable prediction. Namely, replacing the result of Theorem 1 in Eq. (32) produces an affine relation between the beta of the empiricist and \( \beta_{E,n} \):

\[
\beta_n \approx \frac{\delta(1 - R^2)}{R^2 + \delta} + \left[ 1 - \frac{\delta(1 - R^2)}{R^2 + \delta} \right] \beta_{E,n}.
\]

This relation forms the basis of our empirical tests. If we could measure \( \beta_E \) on individual stocks, then a cross-sectional regression would allow us to test whether a distortion exists at all: if \( \delta > 0 \), the intercept is positive and the slope is lower than one.\(^{11}\)

At this stage, the empirical challenge is to obtain time series of expected returns on the market and, most importantly, on individual securities. From these time series we can then estimate \( \beta_{E,n} \) on each stock and run the distortion test associated with Eq. (33). However, the main issue we investigate in this paper is precisely that investors’ expectations are unobservable. The common approach to deal with this problem is to compute expected returns from factor models (e.g., Fama-French factors). A limitation of this approach is that investors likely possess information that typical asset-pricing factors do not capture—we will

\(^{10}\)This is equivalent to abstracting away from effects that can arise from fitting an unconditional model on a conditional one (Jagannathan and Wang, 1996; Lewellen and Nagel, 2006).

\(^{11}\)A positive intercept can also arise if \( \delta \) is negative and larger in absolute value than \( R^2 \). We discuss this (unlikely) possibility below.
never be able to state confidently that we have taken into account all information that could have been relevant for investors. Recently, Martin (2017) and Martin and Wagner (2017) propose extracting this information all at once from option data. Option prices tell us what the market thinks about future returns and thus what the “average investor’s expectations” are. This is the strategy we adopt here and that we explain next. \(^{12}\)

Martin (2017) derives a lower bound on the equity premium using index option prices. Martin and Wagner (2017) extend this approach to compute expected returns on individual stocks, using index and stock option prices. Expected excess returns are derived on a daily basis at the different maturities of traded options: 30, 91, 182 and 365 days. These papers actually provide “bounds” on expected excess returns, as opposed to expected excess returns themselves. However, all we need for our tests are covariances between expected returns (second moments), as opposed to levels (first moments). As in Martin and Wagner (2017), we compute expected excess returns for S&P 500 firms at the individual stock level. We obtain daily equity index prices and return data from CRSP and daily equity index options on the S&P 500 from OptionMetrics. We replicate the approach from Martin and Wagner (2017) and compute three measures of risk-neutral variance, which are then substituted into a parameter-free formula for expected returns on individual stocks. We use the resulting series of expected excess returns to compute \(\beta_{E,n}\) defined in Eq. (32) on individual stocks.

Based on Eq. (33), we then regress empiricist’s betas \(\tilde{\beta}\) onto “expected betas” \(\beta_E\):

\[
\tilde{\beta}_n = a + b\beta_{E,n} + e_n. \quad (34)
\]

Table 1 shows the estimation results. The intercept is strongly statistically significant at all maturities. In particular, Table 1 shows intercepts ranging from 0.67 to 0.74, with \(t\)-stats exceeding ten in all cases (standard errors are bootstrapped and provided in parentheses). The slope coefficients are all lower than one, with \(t\)-stats highly statistically significant (these \(t\)-stats correspond to the null hypothesis that \(b = 1\)). These results strongly suggest a positive distortion in beta estimates. \(^{13}\)

\(^{12}\)In related calculations, Buss and Vilkov (2012) use forward-looking information extracted from option prices to estimate implied market betas. Using these forward-looking betas, they find a monotonically increasing risk-return relation, with a slope close to the historical equity premium. In the context of our model, what Buss and Vilkov (2012) do is to compute the true betas based on the posterior variance of excess returns of the average investor, as in Corollary 1.1. In light of our theory, we interpret the findings of Buss and Vilkov (2012) as an alternative way of testing our result that the CAPM should look stronger if the empiricist uses the correct covariance matrix of excess returns.

\(^{13}\)If \(\delta\) is negative and larger in absolute value than \(R^2\), the intercept of the regression (34) becomes positive again. But this case is unlikely. When \(\delta < -R^2\), it must be that \(a > a_*\), where \(a_* = \lim_{\delta \to -\infty} a = 1 - R^2\). Taking it to the extreme and assuming that \(\delta\) is huge and negative, with \(R^2 \in [0, 0.11]\), we should obtain \(a > 0.9\). This hypothesis is rejected with the numbers from Table 1.
Table 1: Evidence of distortion in beta estimates. Results of the regression specification (34), in which the true betas of the empiricist, $\tilde{\beta}_n$, are regressed cross-sectionally onto the expected betas, $\beta_{E,n}$, at different horizons according to option maturities. Bootstrapped standard errors are provided in parentheses. The columns labeled “Tests” show $t$-stats for the separate null hypotheses $a = 0$ and $b = 1$.

The regression results of Table 1 allow us to perform a “back-of-the-envelope” calculation of $\delta$. Specifically, Table 1 shows that the intercept at the 365-days horizon belongs to the 90% confidence interval $a \in [0.63, 0.87]$. Furthermore, from Eq. (33) we obtain

$$\delta = \frac{aR^2}{1 - a - R^2},$$

which provides a 90% confidence interval for $\delta$. Determining this confidence interval further requires estimating the coefficient of determination from regressing market excess returns on the average investor’s information, $R^2 \equiv \mathbb{V}[\hat{E}[R_{M}]]/\mathbb{V}[R_{M}]$. This coefficient is positive if returns are predictable. Its empirical magnitude is around 10%, depending on the horizon of option prices (Martin, 2017, Table 1). By comparison, in Cochrane (2011, Table 1), it is approximatively 11%, using return-forecasting regressions; accordingly, Figure 2 considers a range from zero to 11% for $R^2$ on the horizontal axis and depicts the resulting 90% confidence interval for $\delta$ in the shaded area.

The plot shows that the distortion can be significant, ranging from 0.3 to 5 if $R^2 = 11\%$. By comparison, the Vasicek (1973) shrinkage proposed in finance textbooks (Bodie et al., 2007; Berk and DeMarzo, 2007) and adopted by practitioners is $\delta = 0.5$ (our 365-day point estimate and $R^2 = 11\%$ jointly imply $\delta = 0.54$). The 90% confidence interval shows that the distortion can in fact be much larger (Levi and Welch (2017) is the only reference we know that advocates for a larger shrinkage). We emphasize that our calculations leave aside relevant sources of variation that can further magnify the distortion. For instance, we have ignored variation arising from differential information across agents (the last term in Eq. 22). Similarly, we have neglected the fact that betas may vary systematically with the market risk premium or with market volatility (Cederburg and O’Doherty, 2016). In that respect, we believe our assessment of the distortion is conservative. The large numbers in Figure 2
thus suggest that the distortion in beta estimates is substantial.

There is another, simpler approach to obtain a rough estimate of the distortion in beta. In the model, dividing the intercept by the slope of the econometrician’s SML in Eq. (30) yields exactly $\delta$. We repeat this exercise in the data: using yearly returns for five beta-sorted portfolios from 1963 to 2017, as well as the market return and risk-free rate, all of which we obtain from Professor French’s website, Panel A of Table 2 reports estimates of average excess returns and betas in columns (a) and (b). These estimated betas, together with estimated average returns, yield the slope and intercept estimates reported in Panel B of Table 2. The failure of the CAPM is evident: the intercept is strongly positive and the slope is significantly lower than the market risk premium, which equals 6.9% in this sample.

Dividing the intercept by the slope, based on the numbers in Panel B of Table 2, yields a distortion of $\delta = 2.96$, which falls within the plausible range plotted in Figure 2.\textsuperscript{14} This (rough) estimate of the distortion, together with Eq. (28), can be used to shrink sample betas towards true betas. Adjusted betas are given in column (c) of Table 2, and imply a CAPM relationship (Panel B) that can no longer be rejected—the intercept is now zero (by construction) and the slope does not differ significantly from 6.9%.

We emphasize that the adjustment applied in column (c) is merely illustrative. In fact,

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Empirically plausible range for the distortion $\delta$. The shaded area shows the 90 percent confidence region for $\delta$ based on Eq. (35) and a 90 percent confidence range for the intercept $a$: $a \in [0.63, 0.87]$. The distortion is plotted as a function of the coefficient of determination from regressing excess returns of the market portfolio on the information set of the average investor, $R^2 \equiv \frac{\text{Var}[E_a[R_M]]}{\text{Var}[R_M]}$.}
\end{figure
Panel A: Five beta-sorted portfolios

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Avg. excess returns (%)</th>
<th>Sample betas</th>
<th>Adj. betas $\delta = 2.96$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low beta</td>
<td>6.90</td>
<td>0.69</td>
<td>0.92</td>
</tr>
<tr>
<td>2</td>
<td>7.85</td>
<td>0.91</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>7.55</td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>7.93</td>
<td>1.22</td>
<td>1.06</td>
</tr>
<tr>
<td>High beta</td>
<td>8.65</td>
<td>1.50</td>
<td>1.13</td>
</tr>
</tbody>
</table>

Panel B: Securities Market Line

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Slope</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.73***</td>
<td>1.93***</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>$-2.4 \times 10^{-13}$</td>
<td>$7.66 ***$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.63)</td>
<td>(1.60)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Observed vs. true SML Columns (a) and (b) of Panel A report average yearly excess returns for five beta-sorted portfolios, using yearly returns from 1963 to 2017, the market return, and the risk-free rate from Professor French’s website. Column (c) adjusts betas according to Eq. (28), for $\delta = 2.96$ (a number obtained by dividing the intercept by the slope). Panel B reports the intercept and the slope of the fitted Securities Market Line in each case. Standard errors of regression estimates are provided in brackets.

since this adjustment involves a linear transformation of betas using only one parameter ($\delta$), it does not change the $R^2$ of the initial relationship, which remains 0.88 both for econometrician’s SML in column (b) and for the “true” SML in column (c). Improving the $R^2$ requires additional degrees of freedom: in the data the distortion is likely characterized by more than just one parameter. For instance, in the model the market proxy is equally weighted, precluding size effects, and payoffs are driven by a single common factor, as opposed to many. Extending the model along these dimensions preserves the validity of the true unconditional CAPM, while allowing the distortion to be characterized by multiple parameters. In conclusion, a more sophisticated model is necessary to take a serious shot at “resurrecting the CAPM” and to truly improve the $R^2$ of the fit.

3 Implications of beta distortion

Several phenomena result from the distorted CAPM relationship of Theorem 1. First, beta distortion offers an alternative interpretation of two well-known anomalies: the underperformance of high-beta stocks (Friend and Blume, 1970; Black et al., 1972), exploited by “betting against beta” (Frazzini and Pedersen, 2014), and the underperformance of high...
idiosyncratic volatility stocks (Ang et al., 2006, 2009). Furthermore, in a dynamic version of this model in which a public announcement is made periodically, the CAPM relation is stronger on announcement days (Savor and Wilson, 2014).

3.1 Reinterpretation of two anomalies

Beta distortion provides a common economic cause of two anomalies. Consider first a long-short portfolio that goes long low-beta stocks, leveraged to a beta of one, and short high-beta stocks, de-leveraged to a beta of one. Frazzini and Pedersen (2014) call this investment strategy “betting against beta” (BAB), and show that it generates positive abnormal returns. The abnormal returns associated with BAB result directly from the observation that the SML is flat (Friend and Blume, 1970; Black et al., 1972) and has a positive intercept. Building on insights from Black (1972), Frazzini and Pedersen (2014) interpret the positive intercept of the SML as a measure of the tightness of investors’ borrowing constraints.

Our interpretation differs. In our setting, betting against beta also earns positive expected excess returns, but for a different reason:

\[
E[R_{BAB}] = \left( \frac{1}{\tilde{\beta}_L} - \frac{1}{\tilde{\beta}_H} \right) \frac{\delta}{1 + \delta} E[R_M] > 0, \tag{36}
\]

where \(\tilde{\beta}_L < 1\) and \(\tilde{\beta}_H > 1\). To obtain Eq. (36), we follow Frazzini and Pedersen (2014) and build a strategy with a long, leveraged position \((1/\tilde{\beta}_L)\) in a low-beta portfolio and a short, de-leveraged position \((-1/\tilde{\beta}_H)\) in a high-beta portfolio. By construction, this strategy has a measured beta of zero. Thus, its returns in excess of the risk-free rate are abnormal. As Eq. (36) shows, if \(\delta > 0\) these abnormal returns are strictly positive.

However, what Eq. (36) also shows is that the abnormal returns earned by betting against beta are in fact reward for taking systematic risk. Because the empiricist overestimates high betas and underestimates low betas, high-beta portfolios seem riskier than they really are and low-beta portfolios seem safer than they really are. From this perspective, betting against measured beta really is betting on true beta.

In a recent paper, Cederburg and O’Doherty (2016) show that the perceived abnormal performance of the BAB strategy disappears after properly incorporating conditioning information. Although this finding suggests that the alpha associated with BAB may result from the informational distance between the empiricist and market participants, it is not definitive evidence for our mechanism and against BAB.\(^\text{15}\) These two mechanisms likely operate

\(^\text{15}\)The BAB strategy analyzed by Cederburg and O’Doherty (2016) is somewhat different from the Frazzini and Pedersen (2014) strategy (it is not beta-neutral).
together in financial markets. We claim that BAB alpha is in fact market risk premium (Eq. 36), a claim that can perhaps be used to distinguish the two theories empirically.

Another immediate consequence of beta distortion is the mis-measurement of idiosyncratic volatility. Because the empiricist mis-estimates beta, her perception of idiosyncratic variance on an individual stock \( n \),

\[
\tilde{\sigma}_{n, id}^2 \equiv \mathbb{V}[R_n] - \tilde{\beta}_n^2 \mathbb{V}[R_M],
\]

is mechanically distorted. In particular, denoting by \( \sigma_{n, id}^2 \) the true idiosyncratic variance, the mis-measurement in idiosyncratic variance depends on the distortion in beta:

\[
\tilde{\sigma}_{n, id}^2 - \sigma_{n, id}^2 = (\beta_n^2 - \tilde{\beta}_n^2) \mathbb{V}[R_M].
\]

Based on the relation between measured and true betas in Theorem 1, the empiricist underestimates idiosyncratic volatility for high-beta stocks and overestimates idiosyncratic volatility for low-beta stocks. Assuming all assets have the same true idiosyncratic variance, the empiricist perceives a negative relation between idiosyncratic volatility and expected returns (the idiosyncratic volatility puzzle). We conclude that, in the context of Theorem 1, betting against beta and the idiosyncratic volatility puzzle are empirical illusions that commonly originate from beta distortion.

### 3.2 CAPM on public announcement days

We now consider a dynamic version of the model of Section 2.2, assuming a public announcement is made periodically (e.g., FOMC meetings, unemployment, or inflation announcements). We show that the CAPM relation measured by the empiricist is stronger on announcement days relative to non-announcement days (Savor and Wilson, 2014). This result follows as a dynamic implication of the distorted view in Eq. (30). Public announcements not only reduce the informational distance between the empiricist and investors, they involve macroeconomic uncertainty that raises the risk premium prior to their release. The empiricist thus observes a steeper SML on announcement days.

Consider a discrete-time economy that goes on forever (e.g. Spiegel, 1998; Watanabe, 2008; Andrei, 2013). The economy is populated with a continuum of investors indexed by \( i \in [0, 1] \), who have CARA utility with common risk aversion \( \gamma \). Each investor \( i \) lives for two periods, entering period \( t \) with wealth \( W_i^t \) and consuming \( W_i^{t+1} \) next period. There are \( N \) risky assets (stocks) and an exogenous riskless bond with constant gross interest rate \( R_f > 1 \). At each period \( t \), the \( N \) stocks pay a vector \( D_t \) of dividends satisfying the dynamic equivalent
of the common factor structure in Eq. (9):

\[ D_t = \bar{D} \mathbf{1} + \Phi F_t + \epsilon_t, \]  

(39)

where \( \epsilon_t \sim \mathcal{N}(0_N, \tau^{-1}_\epsilon I_N) \) is an i.i.d. asset-specific innovation and \( F \) denotes a factor that commonly affects all dividends in the cross section of stocks. As in Section 2.2, the vector \( \Phi \) of loadings is the only source of heterogeneity across assets.

We assume the factor \( F_t \) mean-reverts over time around zero according to

\[ F_t = \kappa_F F_{t-1} + \epsilon^F_t, \quad \epsilon^F_t \sim \text{i.i.d. } \mathcal{N}(0, \tau^{-1}_F), \]  

(40)

with \( 0 \leq \kappa_F \leq 1 \). As in Section 2.2, the market portfolio—the per-capita supply of stocks—is random. We further assume that the market portfolio mean-reverts around its average, \( \bar{M} \), a vector of dimension \( N \) with identical elements that sum up to one, \( \bar{M} = 1/N \):

\[ M_t = (1 - \kappa_M) \bar{M} + \kappa_M M_{t-1} + \epsilon^M_t, \quad \epsilon^M_t \sim \text{i.i.d. } \mathcal{N}(0_N, \tau^{-1}_M I_N), \]  

(41)

with \( 0 \leq \kappa_M \leq 1 \).

Investors observe private and public information. Formally, at any date \( t \), each investor \( i \) receives a private signal \( V^i_t \) about the current factor innovation and a public signal \( G_t \) about the current value of the common factor:

\[ V^i_t = \epsilon^F_t + v^i_t, \quad v^i_t \sim \mathcal{N}(0, \tau^{-1}_v), \quad v^i_t \perp v^j_t, \quad \forall i \neq j \]  

(42)

\[ G_t = F_t + v_t, \quad v_t \sim \mathcal{N}(0, \tau^{-1}_{G,t}). \]  

(43)

The main feature we introduce in this dynamic version of the model is that announcement days (A-day) occur periodically, every \( T \) periods. During an A-day, a pre-scheduled public announcement is released, which increases the precision of public information relative to non-announcement days (N-days):

\[ \tau_{G,k} = \begin{cases} 
\tau_G > \tau_{G,0}^0: & \text{on A-days,} \\
\tau_{G,0}^0: & \text{on N-days.} 
\end{cases} \]  

(44)

Figure 3 illustrates the periodic sequence of events. As a convention, we assume that the announcement takes place at time \( t \). We further assume that all investors commonly observe the fundamental at lag \( T \) and beyond (Townsend, 1983; Singleton, 1987).\(^{16}\)

\(^{16}\)We make this assumption for tractability—without it, the information structure would introduce an infinite-regress inference problem whereby investors would need to infer unobservable shocks over infinitely
Figure 3: Timeline of periodic public announcements days (A-days) in a dynamic model.

We focus on equilibria in which the price $P$ is a linear function of the state variables of the economy. Appendix A.7 provides details on equilibrium derivations. Each investor $i$ builds at time $t - k$ a forecast of cash flows next period, $P_{t-k+1} + D_{t-k+1}$. Since investors are myopic and have CARA preferences, at each lag $k \in \{1, ..., T - 1\}$ in the announcement cycle their asset demands take the following standard form:

$$x^i_{t-k} = \frac{1}{\gamma} \mathbb{E}_{t-k}^{-1} \left[ P_{t-k+1} + D_{t-k+1} \right] \mathbb{E}_{t-k}^i \left[ P_{t-k+1} + D_{t-k+1} - R_f P_{t-k} \right].$$

(45)

In Section 2.2 we have defined the average investor as a fictitious agent who holds consensus beliefs. Keeping the same notation, we denote by $(\hat{\mathbb{E}}_{t-k}[\cdot], \hat{\mathbb{V}}_{t-k}[\cdot])$ the dynamic equivalent to consensus beliefs at lag $k$ in the cycle. In contrast to the static beliefs of Section 2.2, however, these beliefs move over the announcement cycle, hence the index $t - k$.

Because the market portfolio $M_{t-k}$ is mean-variance efficient for the average investor, we obtain a conditional pricing relation at each date of the announcement cycle:

$$\hat{\mathbb{E}}_{t-k} [R_{t-k+1}] = \gamma \hat{\mathbb{V}}_{t-k} [R_{t-k+1}] M_{t-k},$$

(46)

where $R_{t-k+1} \equiv P_{t-k+1} + D_{t-k+1} - R_f P_{t-k}$ represents the dollar excess returns on individual securities. An important difference with the static model of Section 2.2 is that the covariance matrix of returns under the average investor’s information set moves deterministically with each lag $k$. As a result, risk premia now vary over the announcement cycle.

The risk-return tradeoff in Eq. (46) yields a conditional CAPM relation that holds on each day in the announcement cycle. This conditional CAPM is obtained after taking unmany periods back in time. This assumption eliminates the infinite-regress problem: at any time $s$, the fundamental $F_{s-T}$ becomes public information and thus investors only need to infer unobservable shocks up to lag $T - 1$.

With $N$ assets, there are $2^N$ linear equilibria in this model. We focus on the low-volatility equilibrium, as this is the only stable equilibrium to which a finite horizon economy would converge. Bacchetta and Wincoop (2008), Banerjee (2010), Watanabe (2008), and Andrei (2013) discuss the multiplicity of equilibria in infinite horizon models.

---

17With $N$ assets, there are $2^N$ linear equilibria in this model. We focus on the low-volatility equilibrium, as this is the only stable equilibrium to which a finite horizon economy would converge. Bacchetta and Wincoop (2008), Banerjee (2010), Watanabe (2008), and Andrei (2013) discuss the multiplicity of equilibria in infinite horizon models.
conditional expectations at each lag \( k \) for individual assets and for the market:

\[
\mathbb{E}_{t-k}[R_{t-k+1}] = \frac{\widehat{\text{Cov}}[R_{t-k+1}, R_{M,t-k+1}]}{\widehat{\nu}_{t-k}[R_{M,t-k+1}]} \mathbb{E}_{t-k}[R_{M,t-k+1}] = \beta_{t-k} \mathbb{E}_{t-k}[R_{M,t-k+1}],
\]

where the excess return on the market portfolio is defined as \( R_{M,t-k+1} \equiv M'R_{t-k+1} \).

Because risk premia are larger on days during which uncertainty is resolved, the conditional market risk premium is lower on N-days (after public information is revealed) and higher on A-days (before public information is revealed). In Panel (a) of Figure 4, the solid line depicts the conditional CAPM on A-days and the dashed line depicts the conditional CAPM averaged over N-days. This plot is based on an economy with \( N = 4 \) and an A-day every \( T = 4 \) periods.\(^{18}\) Whereas a true CAPM relation holds for both types of days, it is steeper on A-days. Note that averaging over all days yields the true unconditional CAPM.

We now adopt the point of view of the empiricist. Following the steps of Section 2.3, we start with the law of total covariance, which holds at each lag in the announcement cycle:

\[
\nu_{t-k}[R_{t-k+1}] = \widehat{\nu}_{t-k}[R_{t-k+1}] + \nu[\widehat{\nu}_{t-k}[R_{t-k+1}]] + \frac{\Pi_k \Pi_k'}{\tau_v},
\]

where \( \Pi_k \) is a matrix of conformable dimension that multiplies the vector of private signals in each investor’s individual expectation, \( \mathbb{E}_{t-k}[R_{t-k+1}] \). The market-clearing condition (46) then produces a relationship between the variance of excess returns as measured by the empiricist and the variance of excess returns under the information set of the average agent:

\[
\nu_{t-k}[R_{t-k+1}] = \widehat{\nu}_{t-k}[R_{t-k+1}] + \gamma^2 \nu[\widehat{\nu}_{t-k}[R_{t-k+1}]] M_t + \frac{\Pi_k \Pi_k'}{\tau_v}. \tag{49}
\]

The vector of empiricist’s betas is defined analogously to the static setup, except that in this dynamic setup this vector is indexed by \( t - k \):

\[
\tilde{\beta}_{t-k} = \frac{\text{Cov}[R_{t-k+1}, R_{M,t-k+1}]}{\nu_{t-k}[R_{M,t-k+1}]}. \tag{50}
\]

As in the static setup, the informational distance between the empiricist and the average agent distorts the vector of observed betas. Panel (b) of Figure 4 plots the observed conditional CAPM on A-days and on N-days. We follow here the methodology in Savor and Wilson (2014) and let the empiricist estimate betas over the whole sample.\(^{19}\) Using this set

\(^{18}\)Other parameters are: \( \gamma = 2, \tau_F = 275, \tau_M = 120, \tau_v = 1, \phi_1 = 1, \phi_2 = 1/2, \phi_3 = 1/3, \phi_4 = 1/4, \kappa_F = \kappa_M = 0.99, \tau_\epsilon = 9.75, \tau_G^0 = 10^{-6}, \tau_G = 10^3, \bar{M} = 1/4, D = 0, \) and \( R = 1.22. \)

\(^{19}\)Our results do not depend on holding the betas constant. In separate calculations, we estimate different
of unconditional betas, the empiricist observes a steeper SML on A-days, which is also closer to the true SML depicted in panel (a). Furthermore, the CAPM averaged over N-days of the announcement cycle is flatter than the true CAPM depicted in panel (a), with a strong positive intercept. Figure 4 thus provides an illustration whereby the empiricist rejects the CAPM on N-days, but cannot reject it as the correct pricing model on A-days.

Overall, Figure 4 shows that the observed SML becomes steeper on announcement days (Savor and Wilson, 2014), thus moving closer to the true SML. In our model this phenomenon arises through two channels. On the morning of the announcement day, investors forecast how prices will move after the announcement is made. These prices depend on the public signal $G_t$, which investors do not observe yet. Because there is uncertainty about the announcement, the uncertainty about future cash flows, $\hat{V}_{t-1}[P_t + D_t]$, is higher prior to the announcement. Investors thus require a higher risk premium prior to the announcement, increasing the expected return on the market on A-days. Holding betas of assets constant, betas on different types of days. The variation in betas across types of days that we obtain with our calibration is small in magnitude. Thus, the same steepening (flattening) obtains during A-days (N-days).

While we are able to obtain only a numerical characterization of the distortion $\delta$ in this dynamic model (a dynamic model with $T=4$ and $N=4$ has 752 undetermined coefficients), we consistently observe a lower distortion $\delta$ on A-days. This is in line with the intuition that the public announcement is expected to generate systematic price movements (instead of noise) and thus empiricist’s covariance matrix of excess returns should be closer to the true covariance matrix based on investors’ information. It is also in line with the static model: it can be shown that the parameter $\kappa$, which drives the magnitude of the informational distance, decreases with the precision of public information, $\tau_G$. 

Figure 4: CAPM during announcement and non-announcement days. This figure shows the SML during all days (solid), announcement days (dashed), and non-announcement days (dotted). The CAPM looks stronger on announcement days (Savor and Wilson, 2014).
this effect is stronger for high-beta assets (Proposition 1), leading to a steeper SML. Second, the informational distance between investors and the empiricist shrinks as public information is released. Therefore, the distortion in the observed SML is reduced on A-days and for a fleeting moment the observed SML moves closer to the true SML.

The dynamic model developed in this section is merely illustrative, and a full calibration of its parameters to match the data is beyond the scope of this paper. Therefore, it remains an empirical question whether we observe a steeper SML on A-days because the market risk premium is higher or because the informational distance between the empiricist and the average agent shrinks, or both. These two effects take place simultaneously in our model, and disentangling them in the data is a challenging exercise. Recent theoretical papers propose different mechanisms through which the risk premium rises on announcement days (Savor and Wilson, 2013; Ai and Bansal, 2018; Wachter and Zhu, 2018). Whichever mechanism operates in the data, the empiricist’s CAPM in Eq. (30) indifferently implies that an increase in the risk premium leads to a steeper SML on A-days.

4 Conclusion

It has not escaped our notice that the distortion analyzed in this article carries over to factor models in asset pricing. Some variables may appear to the empiricist as priced factors simply because betas are mis-measured. Rather than being priced factors, these variables are instruments for the measurement error in betas. Do there exist economic criteria that would allow the empiricist to distinguish variables that are economically meaningful from those that are not? This matter opens up fascinating directions for future research.

The model can be extended to include other dimensions of heterogeneity among securities. In reality, assets’ payoffs depend on more than one factor, and assets’ market capitalizations are heterogeneous. Extending the model along these dimensions will not affect the validity of the true unconditional CAPM, but will further distort the view of the empiricist. In particular, the empiricist may incorrectly conclude that size, or variables based on valuation ratios, are priced factors.

The basic premise of this paper is that investors actually hold the market portfolio, while the empiricist merely tests whether the market portfolio is mean-variance efficient. Holding the market portfolio presumably gives investors an informational edge over the empiricist. Observing investors’ actions (for instance, their investment decisions) likely reveals some of this information. This approach based on “revealed preference” has caught on recently (Barber et al., 2016; Berk and Van Binsbergen, 2016).
References


A Appendix

A.1 Type I error

This appendix provides detailed derivations for Section 2.1. Use covariance decomposition and apply Stein's lemma to Eq. (2):

\[ 0 = \mathbb{E}[U(W)R|\mathcal{F}] = \mathbb{E}[U(W)(M'R)|\mathcal{F}] \mathbb{E}[R|\mathcal{F}] + \text{Cov}[U(W)(M'R), R|\mathcal{F}] + \mathbb{E}[W(W)(M'R)|\mathcal{F}] \mathbb{V}[R|\mathcal{F}]M, \]

which leads to

\[ \mathbb{E}[R|\mathcal{F}] = -\frac{\mathbb{E}[W(W)(M'R)|\mathcal{F}]}{\mathbb{E}[U(W)(M'R)|\mathcal{F}]} \mathbb{V}[R|\mathcal{F}]M = \gamma \mathbb{V}[R|\mathcal{F}]M. \]  

The second equality (Eq. 3) results from the CARA utility assumption \( U(W) = -e^{-\gamma W} \). Multiply both sides by \( M' \),

\[ \mathbb{E}[R|M] = \gamma \mathbb{V}[R|M]. \]

then divide Eq. (53) by Eq. (54) to obtain the conditional CAPM in this economy (Eq. 4):

\[ \mathbb{E}[R|\mathcal{F}] = \beta_M \mathbb{E}[R|M]. \]

Eq. (5) follows from taking unconditional expectation of Eq. (4) and using covariance decomposition. To prove that the additional terms are together zero, write the unconditional version of Eq. (53),

\[ \mathbb{E}[R] = \gamma \mathbb{V}[R|M]. \]

(Under the Gaussian assumption \( \mathbb{V}[R|M] \) is constant). Then multiply by \( M' \) and divide the equation above by the resulting equation to obtain the unconditional CAPM:

\[ \mathbb{E}[R] = \frac{\mathbb{V}[R|M]M}{\mathbb{V}[R|M]} \mathbb{E}[R|M] = \beta \mathbb{E}[R|M], \]

which implies that the additional terms in Eq. (5) are together zero:

\[ \text{Cov}[\beta_M, \mathbb{E}[R|M]] + (\mathbb{E}[\beta_M] \mathbb{E}[R|M] - \beta \mathbb{E}[R|M]) = 0. \]

Alternatively, one can compute

\[ \text{Cov}[\beta_M, \mathbb{V}[R|M]] = -\mathbb{E}[\beta_M^2] \mathbb{E}[\mathbb{V}[R|M]] + \mathbb{E}[\beta_M \mathbb{V}[R|M]] \]

\[ = -\frac{1}{\gamma} \mathbb{E}[\beta_M] \mathbb{E}[R|M] + \mathbb{E}[\mathbb{V}[R|M]M] \frac{\mathbb{V}[R|M]}{\mathbb{V}[R|M]} \mathbb{V}[R|M] \]

\[ = -\frac{1}{\gamma} \mathbb{E}[\beta_M] \mathbb{E}[R|M] + \mathbb{V}[R|M]M \]

\[ = -\frac{1}{\gamma} \mathbb{E}[\beta_M] \mathbb{E}[R|M] + \frac{1}{\gamma} \mathbb{E}[R] \]

31
$= -\frac{1}{\gamma} E[\beta_M] E[R_M] + \frac{1}{\gamma} \beta E[R_M]$, \hspace{1cm} (63)

and thus

$E[\beta_M] E[R_M] - \beta E[R_M] = -\gamma \text{Cov}[\beta_M, V[R_M|\mathcal{F}]]$, \hspace{1cm} (64)

which proves Eq. (6). Then, using the risk-return tradeoff (54), we obtain again (58).

To obtain Eq. (8) (the type I error), use the definition of covariance:

$E[U_W(M'R)R] = E[U_W(M'R)] E[R] + \text{Cov}[U_W(M'R), R]$ \hspace{1cm} (65)

$= E[U_W(M'R)] E[R] + E[U_{WW}(M'R)] V[R|M]$ \hspace{1cm} (66)

$= \left( -\frac{E[U_{WW}(M'R)]}{\gamma} \right) \gamma V[R|M]M + E[U_{WW}(M'R)] V[R|M]$ \hspace{1cm} (67)

$= E[U_{WW}(M'R)](V[R] - V[R|M])M$. \hspace{1cm} (68)

The third equality follows from using the unconditional version of Eq. (3) and from the CARA utility assumption. The law of total covariance shows that the term $V[R] - V[R|M]$ is non-zero:

$V[R] = V[R|\mathcal{F}] + V[E[R|\mathcal{F}]]$ \hspace{1cm} (69)

$= V[R|M] + \gamma V[R|\mathcal{F}]M$ \hspace{1cm} (70)

$= \gamma^2 V[R|M]^2 V[M]$, \hspace{1cm} (71)

and thus, as long as the the market portfolio $M$ is subject to noise,

$V[R] - V[R|M] = \gamma^2 V[R|M]^2 V[M] \neq 0$. \hspace{1cm} (72)

### A.2 The Type I error persists at high frequency

This appendix proves that the argument of Section 2.1 is robust to the frequency at which returns are measured. For this purpose we repeat the steps of Section 2.1 in the idealized case of diffusions.

For the unconditional CAPM to hold in continuous time we need to modify modeling assumptions appropriately. The canonical framework in which it does is an infinite-horizon economy in which the representative investor has logarithmic utility over consumption and payoffs, $D$, are log-normally distributed (Cochrane (2005), Chapter 9.1.4):

$$dD_t = \mu_D D_t dt + \text{diag}(D_t) \Sigma_D dB_t, \hspace{0.5cm} D_0 = \overline{D},$$ \hspace{1cm} (73)

where $B$ is a $N$-dimensional Brownian motion, and $\mu_D$ and $\Sigma_D$ are constant matrices. Note, however, that the typical framework considered in the literature involves a single risky stock. Extending the framework to $N$ stocks while preserving the CAPM requires additional assumptions to ensure aggregate consumption remains log-normally distributed. For instance, the investor receives a private income that is so that aggregate consumption evolves as a Geometric Brownian motion. In this context equilibrium prices $P_t = \Lambda D_t$ are a constant multiple $\Lambda$ (a constant matrix) of dividends. As a result, excess returns, $R$, are normally distributed in equilibrium:

$$dR_t = \text{diag}(P_t)^{-1}(dP_t + \text{diag}(M_t)^{-1}D_t dt) - r1 dt \equiv \Lambda \mu_D dt + \Lambda \Sigma_D dB_t,$$ \hspace{1cm} (74)

where $M$ is the market portfolio and $r$ is the risk-free rate, which is constant in equilibrium. Under
logarithmic utility it is convenient to model the market portfolio in terms of number of shares available, as opposed to the dollar supply. This way dividends $D = \text{diag}(M)\delta$ above represents the dividend $\delta$ paid for each share held times the number of shares available. Rather than modeling $M$ and $\delta$ separately, we choose to model aggregate dividends, $D$, directly. However, it will be important to keep in mind that $M$ and $D$ are related.

We then follow the steps of Section 2.1. Let $U$ be the investor’s value function. The first-order conditions for optimal portfolio and consumption choices yields the standard Euler equation:

$$
\mathbb{E}[dR_t | \mathcal{F}_t] = -\mathbb{E}\left[ \frac{dU_W(W_t)}{U_W(W_t)} dR_t \bigg| \mathcal{F}_t \right] = -\frac{W_tU_W(W_t)}{U_W(W_t)} d\langle R_t \rangle M_t = d\langle R_t \rangle M_t,
$$

where $\mathcal{F}$ represents all asset-pricing relevant information at instant $t$ and $W$ is the investor’s wealth from holding the market $M$. The second equality applies Ito’s lemma to get the pricing kernel and the third equality uses that $U$ depends separately on wealth and is logarithmic in $W$. This equation is the continuous-time counterpart to the Euler equation in Eq. (2).

This conditional CAPM relation conditions down to an unconditional CAPM relation. Take unconditional expectations on the Euler equation and apply the law of iterated expectations:

$$
\mathbb{E}[dR_t] = \mathbb{E}[d\langle R_t \rangle M_t].
$$

Since the quadratic variation of stock returns in equilibrium,

$$
d\langle R_t \rangle = \Lambda \Sigma D (\Lambda \Sigma D)' dt,
$$

is known, the unconditional Euler equation above simplifies to

$$
\mathbb{E}[dR_t] = d\langle R_t \rangle \mathbb{E}[M_t] = d\langle R_t \rangle M \equiv \beta \mathbb{E}\left[ dR_{\overline{M},t} \right],
$$

where $dR_{\overline{M},t} = \overline{M} dR_t$ and $\beta_t = \mathbb{V} \left[ dR_{\overline{M},t} \right]^{-1} d\langle R_{e,t} \rangle$. $\overline{M}$ denotes the vector of unconditional betas. The last equality follows from premultiplying the left-hand side by $\overline{M}'$, using that $\overline{M}' \beta \equiv 1$ and substituting the risk premium back. This relation is the continuous-time counterpart to Eq. (6).

We conclude that the CAPM holds unconditionally in this economy. The question is whether measuring returns at high frequency (continuous time) helps an empiricist reach the same conclusion. We start by rewriting the Euler equation in Eq. (75) as

$$
0 = \mathbb{E}\left[ (dU_W(W_t) + U_W(W_t)) dR_t \bigg| \mathcal{F}_t \right] = \mathbb{E}\left[ dU_W(W_t), R_t \right] + U_W(W_t)dR_t \mid \mathcal{F}_t
$$

The empiricist can then condition this Euler equation down to an unconditional relation:

$$
0 = \mathbb{E}\left[ U_W(W_t) \mathbb{E}\left[ dR_t \mid \mathcal{F}_t \right] \right] + \mathbb{E}\left[ dU_W(W_t), R_t \right],
$$

which is the continuous-time counterpart to Eq. (7). Note, however, that it involves two terms, as opposed to just one in Eq. (7). Both terms measure the extent to which excess returns covary with marginal utility. But the first term measures this covariation for expected returns (drift), whereas the second term measures this covariation for unexpected shocks in returns (diffusion).

As we have emphasized previously, recovering the CAPM from the Euler equation:

$$
0 = \mathbb{E}\left[ U_W(W_t) \left( -\frac{d\langle U_W(W_t), R_t \rangle}{U_W(W_t)} \right) \right] + \mathbb{E}\left[ d\langle U_W(W_t), R_t \rangle \right] = 0
$$

33
requires observing the market portfolio $M$. Since the empiricist does not observe $M$, she needs a proxy for it to carry out the test. Allowing the empiricist to use the correct unconditional proxy $\overline{M}$, the best she can do is to substitute her proxy for aggregate wealth, $\overline{W}$, in Eq. (80):

$$\mathbb{E} \left[ d\langle U_{\overline{W}}(\overline{W}_t), R_t \rangle + U_{\overline{W}}(\overline{W}_t)dR_t \right] = \mathbb{E} \left[ U_{\overline{W}}(\overline{W}_t)d\langle R_t \rangle \overline{M} + U_{\overline{W}}(\overline{W}_t) \left( -\frac{d\langle U_{\overline{W}}(\overline{W}_t), R_t \rangle}{U_{\overline{W}}(\overline{W}_t)} \right) \right]$$

$$= \langle R_t \rangle \mathbb{E} \left[ U_{\overline{W}}(\overline{W}_t) \left( \frac{\overline{W}_t U_{\overline{W}}(\overline{W}_t)}{U_{\overline{W}}(\overline{W}_t)} \overline{M} - \frac{\overline{W}_t U_{\overline{W}}(\overline{W}_t)}{U_{\overline{W}}(\overline{W}_t)} M_t \right) \right]$$

$$= \langle R_t \rangle \text{Cov} \left( M_t, \overline{W}_t^{-1} \right) \neq 0,$$

which represents the continuous-time counterpart to the CAPM rejection in Eq. (8). Since aggregate wealth depends on equilibrium prices, which are a multiple of dividends, and that dividends themselves depend on the number of shares available $M$, the covariance on the last line is different from zero. Although everyone observes quadratic variations of stock returns, not everyone observes how aggregate wealth covaries instantaneously with stock returns, $d\langle W_t, R_t \rangle$. The magnitude of the resulting disagreement between the empiricist and the investor, as measured by the covariance in the last line in Eq. (82), will depend on how we model noise in the market portfolio.

It is important to discard two tempting but incorrect notions. First, comparing Eq. (8) and Eq. (82), it may seem the notion of informational distance disappears in continuous time. Following the static structure of Section 2.1, suppose the economy terminates after a period of length $dt$, which implies that aggregate wealth is $\overline{W} = M'dR$. We then rewrite the first line in Eq. (82) as:

$$0 \neq \mathbb{E} \left[ d\langle U_{\overline{W}}(\overline{M'}dR), R \rangle \right] + \mathbb{E} \left[ U_{\overline{W}}(\overline{M'}dR)dR \right],$$

where we have eliminated time indices. Since $dR$ is Gaussian we can apply Stein’s lemma and repeat the steps in Appendix A.1 to show that:

$$\mathbb{E} \left[ U_{\overline{W}}(\overline{M'}dR)dR \right] = \mathbb{E}[U_{\overline{W}}(\overline{M'}dR)](\overline{M'} - d\langle R \rangle) \overline{M} = 0,$$

which may suggest that the informational distance between the empiricist and the investor vanishes. What actually vanishes is the covariation between proxied marginal utility and expected returns (the first term in Eq. (85)). However, disagreement persists through the covariation between proxied marginal utility and unexpected shocks in returns (the second term in Eq. (85)):

$$\mathbb{E} \left[ d\langle U_{\overline{W}}(\overline{M'}dR), R \rangle \right] = \mathbb{E}[U_{\overline{W}}(\overline{M'}dR)]d\langle R \rangle \overline{M} = O(dt).$$

Hence, the informational distance becomes the quadratic variation in stock returns. Second, since the violation of optimality in Eq. (82) and in the equation above is of order $dt$, it may seem that increasing the measurement frequency can still reduce the distortion in beta, even though it does not eliminate it completely. However, while both instantaneous variance and covariance are of order $dt$, beta is a ratio of the two and thus does not depend on the measurement frequency.
A.3 Proof of Proposition 1

Start by conjecturing a linear price function of the form:

\[
\begin{bmatrix}
P_1 \\
P_2 \\
\vdots \\
P_N \\
\end{bmatrix} = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_N \\
\end{bmatrix} F + \begin{bmatrix}
g_1 \\
g_2 \\
\vdots \\
g_N \\
\end{bmatrix} G + \begin{bmatrix}
\xi_{11} & \xi_{12} & \cdots & \xi_{1N} \\
\xi_{21} & \xi_{22} & \cdots & \xi_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\xi_{N1} & \xi_{N2} & \cdots & \xi_{NN} \\
\end{bmatrix} \begin{bmatrix}
M_1 \\
M_2 \\
\vdots \\
M_N \\
\end{bmatrix}
\]

(88)

where the undetermined coefficients multiplying the random variables \(F, G,\) and \(M\) will be pinned down by the market clearing condition. Any investor \(i\) has three sources of information gathered in \(F_i\) in Eq. (13): (i) \(N\) public prices, (ii) one private signal \(V_i\), (iii) and one public signal \(G\). We isolate the informational part of prices by subtracting the (known) public signal:

\[
P^a = P - gG = \alpha F + \xi M,
\]

(89)

and stack all information of investor \(i\), both private and public, into a single vector

\[
S^i = \begin{bmatrix}
P^a \\
V^i \\
G \\
\end{bmatrix} = \begin{bmatrix}
\alpha \\
1 \\
1 \\
\end{bmatrix} F + \begin{bmatrix}
\xi & 0_{N \times 1} & 0_{N \times 1} \\
0_{1 \times N} & 1 & 0 \\
0_{1 \times N} & 0 & 1 \\
\end{bmatrix} \begin{bmatrix}
M \\
v^i \\
\end{bmatrix} \equiv HF + \Theta \begin{bmatrix}
M \\
v^i \\
\end{bmatrix},
\]

(90)

where the vector of noise in the signals, \([M v^i v^i]^t\), is jointly Gaussian with covariance matrix:

\[
\Sigma = \begin{bmatrix}
\tau_M^{-1}I_N & 0_{N \times 1} & 0_{N \times 1} \\
0_{1 \times N} & \tau_v^{-1} & 0 \\
0_{1 \times N} & 0 & \tau_G^{-1} \\
\end{bmatrix}.
\]

(91)

Applying standard projection techniques we define

\[
r \equiv (\Theta \Sigma \Theta^t)^{-1} = \begin{bmatrix}
\tau_M (\xi \xi'^t)^{-1} & 0_{N \times 1} & 0_{N \times 1} \\
0_{1 \times N} & \tau_v & 0 \\
0_{1 \times N} & 0 & \tau_G \\
\end{bmatrix},
\]

(92)

and obtain that an investor \(i\)'s total precision on the common factor satisfies

\[
\tau \equiv (\mathbb{V}[F|\mathcal{F}^i])^{-1} = \tau_F + H' r H = \tau_F + \tau_G + \tau_v + \tau_M \alpha' (\xi \xi'^t)^{-1} \alpha.
\]

(93)

The precision \(\tau\) is the same across investors. Furthermore, an investor \(i\)'s expectation of \(F\) satisfies

\[
\mathbb{E}[F|\mathcal{F}^i] = \frac{1}{\tau} H' r S^i = \frac{1}{\tau} \begin{bmatrix}
\alpha' (\xi \xi'^t)^{-1} \tau_M & \tau_v & \tau_G \\
\end{bmatrix} S^i.
\]

(94)

Using the definition of the total precision (93), it follows that average market expectation regarding dividends is

\[
\mathbb{E}[D] = \Phi \frac{1}{\tau} \left[ (\tau - \tau_F - \tau_G) F + \tau_G G + \tau_M \alpha' (\xi \xi'^t)^{-1} \xi M \right] = \mathbb{E}[D|\mathcal{F}^i] - \Phi \frac{\tau_v}{\tau} v^i,
\]

(95)
and average market uncertainty regarding dividends is

$$\hat{V}[D] = \frac{1}{\tau} \Phi\Phi' + \frac{1}{\tau_{\epsilon}} \text{I}_N.$$  \hspace{1cm} (96)

Because agents hold mean-variance portfolios, the market-clearing condition implies:

$$P = \hat{\mathbb{E}}[D] - \gamma \hat{V}[D] M = \Phi \frac{\tau - \tau_F - \tau_G}{\tau} F + \Phi \frac{\tau_G}{\tau} G + \left[ \Phi \frac{\tau_M}{\tau} (\xi^{-1} \alpha)' - \gamma \left( \frac{1}{\tau} \Phi\Phi' + \frac{1}{\tau_{\epsilon}} \text{I}_N \right) \right] M,$$  \hspace{1cm} (97)

where we have used the simplification

$$\tau_M \alpha' (\xi \xi')^{-1} \xi = \tau_M (\xi^{-1} \alpha)'.$$  \hspace{1cm} (98)

The initial price conjecture then yields the following fixed point:

$$\alpha = \Phi \frac{\tau - \tau_F - \tau_G}{\tau},$$  \hspace{1cm} (99)

$$g = \Phi \frac{\tau_G}{\tau},$$  \hspace{1cm} (100)

$$\xi = \Phi \frac{\tau_M}{\tau} (\xi^{-1} \alpha)' - \gamma \left( \frac{1}{\tau} \Phi\Phi' + \frac{1}{\tau_{\epsilon}} \text{I}_N \right).$$  \hspace{1cm} (101)

Multiply both sides of Eq. (101) by $\xi^{-1} \alpha$ (to the right):

$$\alpha = \Phi \frac{\tau_M}{\tau} (\xi^{-1} \alpha)' \xi^{-1} \alpha - \gamma \left( \frac{1}{\tau} \Phi\Phi' + \frac{1}{\tau_{\epsilon}} \text{I}_N \right) \xi^{-1} \alpha,$$  \hspace{1cm} (102)

and then recognize that $\tau_M (\xi^{-1} \alpha)' \xi^{-1} \alpha = \tau_M \alpha' (\xi \xi')^{-1} \alpha = \tau - \tau_F - \tau_G - \tau_v$ (from Eq. 93), which can be replaced above, together with the solution for $\alpha$ to obtain (after multiplication with $\tau$):

$$\Phi \tau_v = -\frac{\gamma}{\tau} \left( \Phi\Phi' + \frac{\tau}{\tau_{\epsilon}} \text{I}_N \right) \xi^{-1} \alpha,$$  \hspace{1cm} (103)

which leads to an equation for $\xi^{-1} \alpha$:

$$\xi^{-1} \alpha = -\frac{\tau_v}{\gamma} \left( \Phi\Phi' + \frac{\tau}{\tau_{\epsilon}} \text{I}_N \right)^{-1} \Phi.$$  \hspace{1cm} (104)

Multiply both sides with $\Phi'$ (to the left):

$$\Phi' \xi^{-1} \alpha = -\frac{\tau_v}{\gamma} \Phi' \left( \Phi\Phi' + \frac{\tau}{\tau_{\epsilon}} \text{I}_N \right)^{-1} \Phi = -\frac{\tau_v}{\gamma} \frac{\Phi'}{\text{I}_N} \frac{\Phi'}{\Phi},$$  \hspace{1cm} (105)

where the second equality follows from the Woodbury matrix identity. Conjecture

$$\xi^{-1}\alpha \equiv -\frac{\sqrt{\tau_p}}{\sqrt{\tau_M}} \Phi,$$  \hspace{1cm} (106)
where $\tau_P$ is an unknown positive scalar. Replacing Eq. (106) in Eq. (93) yields the total precision $\tau$ as a function of this scalar:

$$\tau = \tau_F + \tau_G + \tau_v + \tau_P \Phi' \Phi. \quad (107)$$

Furthermore, replacing the conjecture (106) in Eq. (105) yields

$$\frac{\sqrt{\tau_P}}{\sqrt{\tau_M}} = \frac{\tau_v \tau_\epsilon}{\gamma (\tau + \tau_\epsilon \Phi' \Phi)} \quad (108)$$

which leads to Eq. (16) in the text:

$$\tau_P \left[ \tau_F + \tau_v + \tau_G + (\tau_P + \tau_\epsilon) \Phi' \Phi \right] = \frac{\tau_M \tau_\epsilon^2 \tau_v}{\gamma^2}. \quad (109)$$

This is a cubic equation in $\tau_P$. Its discriminant is strictly negative and thus it has a unique real root. Since it cannot have a negative root (the right hand side is strictly positive), it follows that $\tau_P$ is indeed a unique positive scalar, as conjectured. Finally, the conjecture (106) can now be replaced in the fixed point solution (101) to obtain the undetermined coefficients $\xi$:

$$\xi = - \left( \frac{\gamma + \sqrt{\tau_M \tau_P}}{\tau} \Phi \Phi' + \frac{\gamma}{\tau_\epsilon} I_N \right). \quad (110)$$

This completes the proof of Proposition 1.

### A.4 Proof of Lemma 1

Write the excess returns as

$$D - P = \frac{\tau_F}{\tau} \Phi F - \frac{\tau_G}{\tau} \Phi v + \left( \frac{\gamma + \sqrt{\tau_M \tau_P}}{\tau} \Phi \Phi' + \frac{\gamma}{\tau_\epsilon} I_N \right) M + \epsilon, \quad (111)$$

and thus their unconditional variance is

$$\mathbb{V}[D - P] = \left[ \frac{\tau_F + \tau_G + \tau_P \Phi' \Phi}{\tau^2} + \frac{\gamma^2 \Phi' \Phi}{\tau_M \tau_\epsilon \tau} + \frac{2 \gamma (\gamma + \sqrt{\tau_M \tau_P})}{\tau_M \tau_\epsilon \tau} \right] \Phi \Phi' + \left( \frac{1}{\tau_\epsilon} + \frac{\gamma^2}{\tau_M \tau_\epsilon^2} \right) I_N. \quad (112)$$

Develop the term in square brackets:

$$X = \frac{\tau_F + \tau_G + \tau_P \Phi' \Phi}{\tau^2} + \frac{\gamma^2 \Phi' \Phi}{\tau_M \tau_\epsilon \tau} + \frac{2 \gamma^2}{\tau_M \tau_\epsilon \tau} + \frac{2 \gamma \sqrt{\tau_M \tau_P} \tau + \tau_\epsilon \Phi' \Phi}{\tau_\epsilon \tau_M \tau_\epsilon \tau}. \quad (113)$$

We know from Eq. (109) that

$$\gamma \sqrt{\tau_M \tau_P} (\tau + \tau_\epsilon \Phi' \Phi) = \tau_M \tau_\epsilon \tau_v. \quad (114)$$

Replacing this in Eq. (113) yields

$$X = \left( 1 + \frac{2 \gamma^2}{\tau_M \tau_\epsilon} + \frac{\gamma^2 \Phi' \Phi + \tau_M \tau_v}{\tau_M \tau} \right) \frac{1}{\tau}, \quad (115)$$

37
and thus
\[
\mathbb{V}[R] = \left(1 + \frac{2\gamma^2}{\tau_M \tau_e} + \frac{\gamma^2 \Phi' \Phi + \tau_M \tau_v}{\tau_M \tau} \right) \frac{1}{\tau} \Phi \Phi' + \left(1 + \frac{\gamma^2}{\tau_M \tau_e} \right) \frac{1}{\tau_e} \mathbf{I}_N. \tag{116}
\]

Using Eq. (96), it follows that
\[
\mathbb{V}[R] = \left(1 + \frac{2\gamma^2}{\tau_M \tau_e} + \frac{\gamma^2 \Phi' \Phi + \tau_M \tau_v}{\tau_M \tau} \right) \hat{\mathbb{V}}[R] - \kappa \frac{1}{\tau_e} \mathbf{I}_N, \tag{117}
\]
where \( \kappa \) is defined as in (24):
\[
\kappa \equiv \frac{\gamma^2}{\tau_M} \left( \frac{1}{\tau_e} + \frac{\Phi' \Phi}{\tau} \right) + \frac{\tau_v}{\tau}. \tag{118}
\]
Thus,
\[
\mathbb{V}[R] - \hat{\mathbb{V}}[R] = \left(\kappa + \frac{\gamma^2}{\tau_M \tau_e} \right) \hat{\mathbb{V}}[R] - \kappa \frac{1}{\tau_e} \mathbf{I}_N, \tag{119}
\]
Using Eq. (96) again, obtain:
\[
\frac{1}{\tau_e} \mathbf{I}_N = \hat{\mathbb{V}}[R] - \frac{1}{\tau} \Phi \Phi', \tag{120}
\]
which can be replaced in Eq. (119):
\[
\mathbb{V}[R] - \hat{\mathbb{V}}[R] = \left(\kappa + \frac{\gamma^2}{\tau_M \tau_e} \right) \hat{\mathbb{V}}[R] - \kappa \left( \frac{1}{\tau} \Phi \Phi' \right) = \frac{\gamma^2}{\tau_M \tau_e} \hat{\mathbb{V}}[R] + \frac{\kappa}{\tau} \Phi \Phi'. \tag{121}
\]
This completes the proof of Lemma 1.

**A.5 Proof of Theorem 1**

Start with the unconditional true CAPM (20):
\[
\mathbb{E}[R] = \frac{\frac{1}{N} \hat{\mathbb{V}}[R] \mathbf{1}}{\mathbb{V}[R]} \mathbb{E}[R_M], \tag{122}
\]
and replace the following relationship, which results from Eq. (119):
\[
\hat{\mathbb{V}}[R] = \frac{\mathbb{V}[R] + \frac{\kappa}{\tau_e} \mathbf{I}_N}{1 + \kappa + \frac{\gamma^2}{\tau_M \tau_e}}. \tag{123}
\]
This yields
\[
\mathbb{E}[R] = \frac{\frac{\mathbb{V}[R_M]}{1 + \kappa + \frac{\gamma^2}{\tau_M \tau_e}} \tilde{\beta} \mathbb{E}[R_M] + \frac{\kappa}{\tau_e} \frac{1}{N} \frac{\mathbb{V}[R_M]}{1 + \kappa + \frac{\gamma^2}{\tau_M \tau_e}} \hat{\mathbb{V}}[R_M]}{1 + \kappa + \frac{\gamma^2}{\tau_M \tau_e}}, \tag{124}
\]
38
where \( \tilde{\beta} \) is defined in (26). Multiply both sides with the equally weighted market portfolio \( \frac{1}{N} \) and use the fact that the weighted average of empiricist’s betas is one, then divide by \( \mathbb{E}[R_M] \):

\[
1 = \frac{\mathbb{V}[R_M]}{(1 + \kappa + \frac{\gamma^2}{\tau M\tau v}) \mathbb{V}[R_M]} + \frac{\kappa / (\tau e N)}{(1 + \kappa + \frac{\gamma^2}{\tau M\tau v}) \mathbb{V}[R_M]}. \tag{125}
\]

Both terms on the right hand side are positive. Thus, there exists a unique positive scalar \( \delta \) such that the first term equals \( \frac{1}{1 + \delta} \) and the second term equals \( \delta / (1 + \delta) \). This yields

\[
\delta = \frac{\kappa / N}{\tau e \mathbb{V}[R_M]}, \tag{126}
\]

and Eq. (124) can now be written for any individual stock \( n \):

\[
\mathbb{E}[R_n] = \frac{\delta}{1 + \delta} \mathbb{E}[R_M] + \frac{1}{1 + \delta} \gamma_n \mathbb{E}[R_M] = \frac{\delta}{1 + \delta} (1 - \tilde{\beta}_n) \mathbb{E}[R_M] + \tilde{\beta}_n \mathbb{E}[R_M], \tag{127}
\]

which is Eq. (27) in the text. Also, the first equality above is Eq. (30) in the text. The remainder of Theorem 1 follows from recognizing that the empiricist and the average agent obtain the same average expected returns for any asset \( n \in \{1, \ldots, N\} \):

\[
\frac{\delta + \tilde{\beta}_n}{1 + \delta} \mathbb{E}[R_M] = \beta_n \mathbb{E}[R_M], \tag{128}
\]

and thus this relationship between \( \tilde{\beta}_n \) and \( \beta_n \) completes the proof of Theorem 1:

\[
\tilde{\beta}_n - 1 = (1 + \delta)(\beta_n - 1), \quad \forall n \in \{1, \ldots, N\}. \tag{129}
\]

### A.6 Conditioning on public information

The main result of the paper (Theorem 1) assumes the empiricist’s information is limited to realized returns. In this appendix, we augment empiricist’s dataset with all relevant public information.

First, suppose we allow the empiricist to condition on the public signal \( G \) when computing betas of securities. Proposition 1 implies

\[
\text{Cov}[R, G] = 0_{N \times 1}. \tag{130}
\]

Thus, controlling for the public signal cannot improve the empiricist’s measurement (because prices already incorporate all available public information).

Second, suppose the empiricist controls jointly for prices and the signal \( G \). Under this augmented information set, we show below that Theorem 1 still holds, but with a different coefficient of distortion, \( \tilde{\delta} \):

\[
\tilde{\delta} = \frac{1}{N \mathbb{V}[R_M|\{P, G\}]} \frac{\tau v}{\tau v (\tau - \tau v)} \geq 0, \tag{131}
\]

where \( \mathbb{V}[R_M|\{P, G\}] \) denotes the variance of excess returns on the market portfolio conditional on

---

\(^{21}\) Although we have shown above that conditioning on the public signal \( G \) alone does not help, it must be included in the information set when conditioning on prices. If not, then it creates an omitted variable bias (because both prices and realized returns depend on the public signal).
observing all publicly available information:
\[
\mathbb{V}[R]\{P,G\} = \mathbb{V}[D]\{P,G\} = \frac{1}{\tau - \tau_v} \Phi \Phi' + \frac{1}{\tau_\epsilon} \mathbf{I}_N, \tag{132}
\]
and thus we can write\(^\text{22}\)
\[
\mathbb{V}[R]\{P,G\} = \frac{\tau}{\tau - \tau_v} \hat{\mathbb{V}}[R] - \frac{\tau_v}{\tau_\epsilon (\tau - \tau_v)} \mathbf{I}_N. \tag{133}
\]
The empiricist obtains a new set of betas:
\[
\tilde{\beta} = \frac{1}{N} \mathbb{V}[R]\{P,G\} \frac{\mathbf{1}}{\mathbb{V}[R_M]\{P,G\}} = \frac{\tau}{\tau - \tau_v} \hat{\mathbb{V}}[R_M] \beta - \frac{\tau_v}{N \mathbb{V}[R_M]\{P,G\}} \frac{\mathbf{1}}{\tau_\epsilon (\tau - \tau_v)}. \tag{134}
\]
Take average on both sides by multiplying with \(1/N\):
\[
1 = (1 + \tilde{\delta}) - \frac{1}{N \mathbb{V}[R_M]\{P,G\}} \frac{\tau_v}{\tau_\epsilon (\tau - \tau_v)}, \tag{135}
\]
and thus we obtain Eq. (131). Replacing \(\tilde{\delta}\) in (134) and subtracting \(1\) on both sides yields
\[
\tilde{\beta} - 1 = (1 + \tilde{\delta}) (\beta - 1). \tag{136}
\]
Thus, the empiricist still observes distorted betas, even after controlling for all the available public information. Furthermore, it can be shown (through numerical examples) that the distortion \(\tilde{\delta}\) may become larger than the initial distortion obtained without conditioning, \(\delta\). This happens when the private information is sufficiently precise. Thus, the presence of dispersed private information makes the CAPM untestable in principle: because the empiricist cannot control for all the available private information, it will always remain the case that she observes a flatter SML, even after controlling for all publicly available information.

### A.7 Dynamic model

This appendix solves the equilibrium in the dynamic model of Section 3.2. The vector of dividends \(D_t\), the factor \(F_t\) and the random supply follow the processes:
\[
D_t = \mathbf{d} \mathbf{1}_N + \Phi F_t + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.} \ \mathcal{N}(\mathbf{0}_N, \tau_\epsilon^{-1} \mathbf{I}_N) \tag{137}
\]
\[
F_t = \kappa_F F_{t-1} + \epsilon_t^F, \quad \epsilon_t^F \sim \text{i.i.d.} \ \mathcal{N}(0, \tau_F^{-1}) \tag{138}
\]
\[
M_t = (1 - \kappa_M) \bar{M} + \kappa_M M_{t-1} + \epsilon_t^M, \quad \epsilon_t^M \sim \text{i.i.d.} \ \mathcal{N}(\mathbf{0}_N, \tau_M^{-1} \mathbf{I}_N), \tag{139}
\]
where \(0 \leq \kappa_F \leq 1, 0 \leq \kappa_M \leq 1\), and \(\bar{M} = \mathbf{1}_N / N\).

\(^{22}\)Notice that it is not necessary to assume here that the empiricist knows the price coefficients. This is because \(\mathbb{V}[R]\{P,G\} = \mathbb{V}[R] - \mathbb{V}[E[R]\{P,G\}]\). The empiricist can compute both terms on the right hand side: the first term is the total covariance matrix; the second term represents explained variation (i.e., the numerator of the coefficient of determination \(R^2\)) after regressing realized returns on prices \(P\) and the public signal \(G\).
Investors receive private and public signals that are informative about the common factor:

\[ V_t^i = \epsilon_t^F + v_t^i, \quad v_t^i \sim \mathcal{N}(0, \tau_v^{-1}), \quad v_t^i \perp v_t^j, \quad \forall i \neq j \tag{140} \]
\[ G_t = F_t + v_t, \quad v_t \sim \mathcal{N}(0, \tau_G^{-1}). \tag{141} \]

The precision of the public signal changes on announcement days, which occur every \( T \) periods:

\[ \tau_{G,k} = \begin{cases} \tau_G > \tau_G^0, & \text{on A-days (i.e., at } t - T \text{ and at } t) \\ \tau_G^0, & \text{on N-days.} \end{cases} \tag{142} \]

We solve simultaneously for all equilibrium prices over the announcement cycle. We use the index \( k \) to model lags in the announcement cycle:

\[ k \in \{1, ..., T\}, \tag{143} \]

and we make the convention that the public announcement is made at \( k = T \). Thus, the A-day is the period between \( t - 1 \) and \( t \), right before the public announcement takes place.

We conjecture that prices at time \( t - k \) take the following linear form:

\[ P_{t-k} = \bar{\alpha}_k \bar{D} + \alpha_k F_{t-k-T} + \xi_k \bar{M} + \xi_k M_{t-k-T} + g_k \tilde{G}_{t-k} + a_k \epsilon_{t-k}^F + d_k \tilde{D}_{t-k} + b_k \tilde{\epsilon}_{t-k}^M, \tag{144} \]

where the price coefficients are vectors/matrices,

\[ \bar{\alpha}_k, \alpha_k, \xi_k \in \mathbb{R}^{N \times 1}, \xi_k \in \mathbb{R}^{N \times N}, g_k, a_k \in \mathbb{R}^{N \times T}, d_k, b_k \in \mathbb{R}^{N \times NT}, \tag{145} \]

and we use the following notation:

\[
\begin{align*}
\bar{G}_{t-k} &\equiv \begin{bmatrix} G_{t-k-T+1} \\ G_{t-k-T+2} \\ \vdots \\ G_{t-k} \end{bmatrix}, \quad \epsilon_{t-k}^F &\equiv \begin{bmatrix} \epsilon_{t-k-T+1}^F \\ \epsilon_{t-k-T+2}^F \\ \vdots \\ \epsilon_{t-k}^F \end{bmatrix}, \quad \bar{D}_{t-k} &\equiv \begin{bmatrix} D_{t-k-T+1} \\ D_{t-k-T+2} \\ \vdots \\ D_{t-k} \end{bmatrix}, \quad \tilde{\epsilon}_{t-k}^M &\equiv \begin{bmatrix} \epsilon_{t-k-T+1}^M \\ \epsilon_{t-k-T+2}^M \\ \vdots \\ \epsilon_{t-k}^M \end{bmatrix}.
\end{align*}
\tag{146} \]

More precisely:

- \( \bar{G}_{t-k} \) is a vector of dimension \( T \) containing all the public signals \( G \) from time \( t - k - T + 1 \) to time \( t - k \),
- \( \epsilon_{t-k}^F \) is a vector of dimension \( T \) containing all the fundamental shocks \( \epsilon^F \) from time \( t - k - T + 1 \) to time \( t - k \),
- \( \bar{D}_{t-k} \) is a stacked vector of dimension \( NT \) containing the payoffs of all \( N \) assets from time \( t - k - T + 1 \) to time \( t - k \),
- \( \tilde{\epsilon}_{t-k}^M \) is a stacked vector of dimension \( NT \) containing the supply shocks for all assets from time \( t - k - T + 1 \) to time \( t - k \).

The total number of undetermined coefficients at each lag \( k \) is \( N [3 + 2T + N(1 + 2T)] \). Because coefficients change over the announcement cycle, finding the equilibrium requires solving simultaneously for a total of \( NT [3 + 2T + N(1 + 2T)] \) undetermined coefficients.

To understand the price structure (144), notice that any investor \( i \) observes at date \( t - k \):

\[
\mathcal{F}_t^i \equiv \{ \{ F_s \}_{s \leq t-k-T}, \{ P_s \}_{s \leq t-k}, \{ V_s^i \}_{s \leq t-k}, \{ D_s \}_{s \leq t-k}, \{ G_s \}_{s \leq t-k} \}. \tag{147} \]
It follows that an investor $i$ observes $\{\varepsilon^F_t \}_{s \leq t - k - T}$, but not the vector of innovations $\epsilon^F_{t-k}$, which is only partially revealed by equilibrium prices. Furthermore, the sequence of past dividends $\{D_{t-k-s}\}_{s=0}^{T-1}$ reveals information regarding past, unobservable factor innovations $\{\epsilon^F_{t-k-s}\}_{s=0}^{T-1}$, hence the vector $\mathcal{D}_{t-k}$. The same holds for the vector $\mathcal{G}_{t-k}$. Finally, to understand the structure of the vector of past supply innovations, consider the price at time $t - k - T$. This price reveals a linear combination of $\{\epsilon^F_{t-k-s}\}_{s=0}^{T-1}$ and $M_{t-k-T}$. Since these factor innovations are observable, so is $M_{t-k-T}$, hence the vector $\epsilon^M_{t-k}$ and the presence of $F_{t-k-T}$ and $M_{t-k-T}$ in the equilibrium price.

At time $t - k$, an investor $i$’s information set contains four sources of information regarding the vector $\epsilon^F_{t-k}$: past and current prices, private signals, dividends, and public announcements. We isolate the informational part of prices $P^a_{t-j}$ that only contains unobservables at $t - k$:

$$P^a_{t-j} = \tilde{a}_j \epsilon^F_{t-k} + \tilde{b}_j \epsilon^M_{t-k}, \quad \text{for } j \in \{1, \ldots, T\}$$

which we stack into a single vector $\mathbb{P}_{t-k}$ of dimension $NT$:

$$\mathbb{P}_{t-k} \equiv \begin{bmatrix} P^a_{t-k} & P^a_{t-k-1} & \cdots & P^a_{t-k-T+2} & P^a_{t-k-T+1} \end{bmatrix}'.$$

We can further write

$$\mathbb{P}_{t-k} = \mathbb{A}_{k} \epsilon^F_{t-k} + \mathbb{B}_{k} \epsilon^M_{t-k},$$

where we construct the matrices $\mathbb{A}_{k} \in \mathbb{R}^{NT \times T}$ and $\mathbb{B}_{k} \in \mathbb{R}^{NT \times NT}$ from $\tilde{a}_j$ and $\tilde{b}_j$, $j \in \{1, \ldots, T\}$. Similarly, we collect all past and current private signals that investor $i$ has gathered by time $t - k$ into a vector $\mathcal{V}^i_{t-k}$ of dimension $T$:

$$\mathcal{V}^i_{t-k} \equiv \begin{bmatrix} V^i_{t-k} & V^i_{t-k-1} & \cdots & V^i_{t-k-T+1} \end{bmatrix}'.$$

which can be further written

$$\mathcal{V}^i_{t-k} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \epsilon^F_{t-k} + \begin{bmatrix} v^i_{t-k} \\ v^i_{t-k-1} \\ \vdots \\ v^i_{t-k-T+2} \\ v^i_{t-k-T+1} \end{bmatrix} = \begin{bmatrix} \Omega \epsilon^F_{t-k} \\ \tilde{\epsilon}^i_{t-k} \end{bmatrix}.$$  

Furthermore, dividends also reveal a combination of observable and unobservable information. In particular, the vector $\mathcal{D}_{t-k}$ can be written

$$\mathcal{D}_{t-k} = \mathcal{D} \mathbf{1}_{NT} + \mathcal{A}_{0,t-k} F_{t-k-T} + \mathcal{A}_{t-k} \epsilon^F_{t} + \tilde{\epsilon}_{t-k},$$

where, denoting the Kronecker product by $\otimes$,

$$\mathcal{A}_{0,t-k} \equiv \begin{bmatrix} \kappa^F_{T} \\ \kappa^F_{T-1} \\ \vdots \\ \kappa^F_{0} \end{bmatrix} \otimes \mathcal{G}_{0}, \quad \mathcal{A}_{t-k} \equiv \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \kappa^F_{T} & 1 & 0 & \cdots & 0 \\ \vdots \\ \kappa^F_{T-1} & \kappa^F_{T-2} & \kappa^F_{T-3} & \cdots & 1 \end{bmatrix} \otimes \mathcal{G}, \quad \tilde{\epsilon}_{t-k} \equiv \begin{bmatrix} \epsilon_{t-k-T+1} \\ \epsilon_{t-k-T+2} \\ \vdots \\ \epsilon_{t-k} \end{bmatrix}.$$  

42
Consider the informational part of dividends that only contains unobservables, which we stack into a single vector $\mathbb{D}_t-k$ of dimension $NT$:

$$\mathbb{D}_t-k \equiv \Lambda_{t-k} \epsilon_{t-k}^F + \tilde{\epsilon}_{t-k}. \quad (155)$$

Finally, public announcements have the following structure:

$$\tilde{G}_{t-k} = \Omega_{G0} F_{t-k-T} + \Omega_G F_{t-k} \epsilon_{t-k}^F + \tilde{\epsilon}_{t-k} = \tilde{\epsilon}_{t-k}$$

and thus we can isolate the informational part:

$$G_{t-k} = \Omega_G \epsilon_{t-k}^F + \tilde{\epsilon}_{t-k}. \quad (157)$$

Overall, an investor $i$'s information at time $t-k$ is fully summarized by $P_{t-k}$, $\tilde{V}_i^i$, $\mathbb{D}_t-k$, and $G_{t-k}$. Since the vector grouping this information and the vector $\epsilon_{t-k}^F$ are jointly normally distributed, investor $i$ forms her conditional expectations $E_{i[t-k]}[\epsilon_{t-k}^F]$ and conditional variance $V_{i[t-k]}[\epsilon_{t-k}^F]$ by projecting the former vector on the latter. We write the conditional expectation and conditional variance in Proposition 2.

**Proposition 2.** At time $t-k$, an investor $i$'s conditional expectation and conditional variance of the fundamental innovations $\epsilon_{t-k}^F$ satisfy

$$E_{i[t-k]}[\epsilon_{t-k}^F] = \left( \frac{1}{\tau_{F}} \mathbb{H}_k \right)^{-1} \left( \frac{1}{\tau_{F}} \mathbb{H}_k \mathbb{H}_k' + \mathbb{Q}_k \right)^{-1} \left[ \begin{array}{c} P_{t-k} \\ V_i^i \\ \mathbb{D}_t-k \\ G_{t-k} \end{array} \right]$$

and

$$V_{i[t-k]}[\epsilon_{t-k}^F] = \frac{1}{\tau_{F}} \left( \mathbb{I}_{T} - \mathbb{M}_k \mathbb{H}_k \right)$$

where the matrices $\mathbb{H}_k$ and $\mathbb{Q}_k$ are defined below.

**Proof.** Stack all the information available to agent $i$ at time $t-k$ into a single vector:

$$\begin{bmatrix} P_{t-k} \\ V_i^i \\ \mathbb{D}_t-k \\ G_{t-k} \end{bmatrix} = \begin{bmatrix} \Lambda_{t-k} \\ \mathbb{H} \end{bmatrix} \begin{bmatrix} \epsilon_{t-k}^F \\ \tilde{\epsilon}_{t-k} \end{bmatrix} + \Theta_k \begin{bmatrix} \epsilon_{t-k}^M \\ \tilde{\epsilon}_{t-k} \end{bmatrix}, \quad (160)$$

where $\Theta_k \in \mathbb{R}^{(2NT+2T) \times (2NT+2T)}$ is a block diagonal matrix:

$$\Theta \equiv \text{diag}(\mathbb{B}_k, \mathbb{I}_{NT+2T}). \quad (161)$$
The vector of noise in investor $i$’s information (i.e., the last vector in Eq. (160), denoted by $s_i^{t-k}$) is normally distributed with mean zero and covariance matrix

$$\Sigma_k \equiv \text{diag} \left( \frac{1}{\tau_M} \mathbf{I}_N, \frac{1}{\tau_v} \mathbf{I}_T, \frac{1}{\tau_e} \mathbf{I}_{NT}, \left[ \frac{1}{\tau_{G,k+T-1}} \quad \frac{1}{\tau_{G,k+T-2}} \quad \ldots \quad \frac{1}{\tau_{G,k}} \right] \right).$$  \hspace{1cm} (162)

The last part of this matrix varies over the announcement cycle. Defining $\mathbb{Q}_k \equiv \Theta \Sigma_k \Theta'$, direct application of the projection theorem yields Eqs. (158)--(159).

Each investor $i$’s asset demand takes the following standard form:

$$a_i^t = \frac{\nabla^{i}_{t-k} \left[P_{t-k+1} + D_{t-k+1}\right]}{\gamma} \mathbb{E}^{i}_{t-k} \left[P_{t-k+1} + D_{t-k+1} - R_f P_{t-k}\right].$$  \hspace{1cm} (163)

To write $P_{t-k+1} + D_{t-k+1}$, define first $\mathscr{L}$ as a linear function of $\bar{D}, F_{t-k-T}, \xi, \xi_{t-k-T}, \bar{G}_{t-k}$. The coefficients of this function are obtained directly from writing the price conjecture (144) at period $t - k + 1$ and adding $D_{t-k+1}$. The output of the linear function $\mathscr{L}$ is a vector of dimension $N$:

$$\mathscr{L} = f(\bar{D}, F_{t-k-T}, \xi, \xi_{t-k-T}, \bar{G}_{t-k}).$$  \hspace{1cm} (164)

We can therefore write

$$P_{t-k+1} + D_{t-k+1} = \mathscr{L} + a^*_k e^F_{t-k} + d^*_k \bar{D}_{t-k} + b^*_k e^M_{t-k} + \begin{bmatrix} \tilde{a}_k & \tilde{d}_k & \tilde{b}_k & \tilde{g}_k \end{bmatrix}_{adbg_k} \begin{bmatrix} \varepsilon_{t-k+1} F \\ \varepsilon_{t-k+1} M \\ \tilde{v}_{t-k+1} \\ \tilde{n}_{t-k+1} \end{bmatrix},$$  \hspace{1cm} (165)

where the coefficients

$$a^*_k \in \mathbb{R}^{N \times T}, d^*_k, b^*_k \in \mathbb{R}^{N \times NT}, \tilde{a}_k, \tilde{g}_k \in \mathbb{R}^{N \times 1}, \tilde{d}_k, \tilde{b}_k \in \mathbb{R}^{N \times N}$$  \hspace{1cm} (166)

are obtained directly from writing the price conjecture (144) at period $t - k + 1$ and adding $D_{t-k+1}$. The matrix $adbg_k$ groups together $\tilde{a}_k, \tilde{d}_k, \tilde{b}_k,$ and $\tilde{g}_k$. From Eq. (150), we can write

$$\bar{e}_{t-k}^F = \mathbb{B}_k^{-1} P_{t-k} - \mathbb{B}_k^{-1} A_k e^F_{t-k},$$  \hspace{1cm} (167)

which is further replaced in Eq. (165):

$$P_{t-k+1} + D_{t-k+1} = \mathscr{L} + \left( a^*_k - b^*_k \mathbb{B}_k^{-1} A_k \right) e^F_{t-k} + d^*_k \bar{D}_{t-k} + b^*_k \mathbb{B}_k^{-1} P_{t-k} + adbg_k \tilde{n}_{t-k+1}.$$  \hspace{1cm} (168)

Given this form for future payoffs, we can now use Proposition 2 to compute:

$$\mathbb{E}^{i}_{t-k}[P_{t-k+1} + D_{t-k+1}] = \mathscr{L} + \psi_k \mathbb{M}_k \begin{bmatrix} P_{t-k} \\ V_{t-k}^i \\ D_{t-k} \end{bmatrix}_{\Psi^i_{t-k}} + d^*_k \bar{D}_{t-k} + b^*_k \mathbb{B}_k^{-1} P_{t-k}$$  \hspace{1cm} (169)

$$\nabla^{i}_{t-k}[P_{t-k+1} + D_{t-k+1}] = \psi_k \frac{1}{\tau_F} \left( \mathbb{I}_T - \mathbb{M}_k \mathbb{I}_T \right) \psi'_k + adbg_k \Sigma_n^{k+1} adbg'_k,$$  \hspace{1cm} (170)


44
where $\Sigma_{n,k+1}$ is the covariance matrix of the vector $\tilde{n}_{t-k+1}$:

$$\Sigma_{n,k+1} \equiv diag \left( \frac{1}{\tau_F}, \frac{1}{\tau_e}, \frac{1}{\tau_M}, \frac{1}{\tau_{G,k+1}} \right). \quad (171)$$

This matrix changes over the announcement cycle. Since the precision of private information is homogeneous across agents, $\nabla_{t-k}[P_{t-k+1} + D_{t-k+1}]$ equals the average precision in the economy:

$$\nabla_{t-k}[P_{t-k+1} + D_{t-k+1}] \equiv \nabla_{t-k}^i[P_{t-k+1} + D_{t-k+1}]. \quad (172)$$

Furthermore, the consensus beliefs are defined by averaging over investors’ expectations:

$$\hat{E}_{t-k}[P_{t-k+1} + D_{t-k+1}] \equiv \int_i \hat{E}^i_{t-k}[P_{t-k+1} + D_{t-k+1}] di. \quad (173)$$

The matrix $\Pi_k$ in Eq. (48) is obtained by taking the difference:

$$E^i_{t-k}[P_{t-k+1} + D_{t-k+1}] - \hat{E}_{t-k}[P_{t-k+1} + D_{t-k+1}] = \Pi_k \tilde{v}^i_{t-k}. \quad (174)$$

The market clearing condition at each lag $k$ in the announcement cycle writes:

$$\int_i x^i_{t-k} di = M_{t-k}, \quad \text{for } k \in \{1, \ldots, T\}, \quad (175)$$

and thus the undetermined coefficients solve $T$ simultaneous systems of equations (one for each trading date in the announcement cycle):

$$P_{t-k} = \frac{\hat{E}[P_{t-k+1} + D_{t-k+1}] - \gamma \hat{V}_{t-k}[P_{t-k+1} + D_{t-k+1}] M_{t-k}}{R_f}, \quad \text{for } k \in \{1, \ldots, T\}. \quad (176)$$

where $M_{t-k}$ can be further written

$$M_{t-k} = (1 - \kappa_T^M)M + \kappa_T^M M_{t-k} - T + \left[ \begin{array}{cccc} \kappa_T^{T-1} & \kappa_T^{T-2} & \ldots & \kappa_M \end{array} \right] \otimes I_N \tilde{E}^M_{t-k} \quad (177)$$

The system (176), which has $NT [3 + 2T + N(1 + 2T)]$ equations and $NT [3 + 2T + N(1 + 2T)]$ unknowns, is solved using fixed-point iteration. Using as a starting point the solution from an economy without A-days (in which the number of unknowns is divided by $T$) ensures fast convergence.

With $N$ assets, there are $2^N$ equilibria (Bacchetta and Wincoop, 2008; Banerjee, 2010; Watanabe, 2008; Andrei, 2013). The fixed-point iteration algorithm converges only to the low-volatility equilibrium. This is the only stable equilibrium to which a finite horizon economy would converge. It is the equilibrium that we analyze in Section 3.2.