A Model-Free Assessment of the Importance of Subjective Beliefs for Asset Pricing*

Paymon Khorrami
Duke University

October 18, 2025

Abstract

Are belief dynamics or risks and risk attitudes more important for asset pricing? Allowing both, I use survey data combined with subjective-belief versions of stochastic discount factor (SDF) volatility bounds to shed new light on this classic question. I estimate lower bounds for the volatility of the SDF attributable to (i) risks relevant for investor marginal utility, versus (ii) subjective belief dynamics. The estimates suggest that risks, particularly long-term risks, make up at least 50% of SDF volatility, with many estimates much higher than 50%. An example extrapolation model with a modest direct contribution of beliefs to SDF volatility (about 25%) can account for my estimates. This example also highlights the potential for a novel mechanism, *subjective risk*, which is the indirect impact beliefs have on asset prices through induced marginal utility volatility. Across several models, I show that beliefs must matter primarily through this subjective risk mechanism, because otherwise they create too much transitory SDF volatility.

JEL Codes: G00, G10, G12, G40

Keywords: stochastic discount factor, volatility bounds, subjective beliefs, long-run risk, disagreement

*E-mail: paymon.khorrami@duke.edu. Thanks to Anmol Bhandari, Jarda Borovicka, Adrian Buss, Aditya Chaudhry, Ricardo de la O, Lars Hansen, Spencer Kwon, Sean Myers, Stefan Nagel, and participants at the Yale Junior Finance Workshop, Duke Finance Brownbag, SFS Cavalcade, WFA, EFA, Purdue Daniels, and UChicago Blue Collar Working Group for comments.

1 Introduction

Are beliefs or risks more important for asset pricing? Despite a long line of research, answers remain elusive. A common approach is to write down a fully-specified model (behavioral or rational) and test it against a set of important asset market data. But the results from such models can often depend on assumptions that are hard to verify or measure. For example, while a subjective belief model may account for many properties of asset prices, its success hinges not only on its belief dynamics but also on auxiliary assumptions on preferences, frictions, and heterogeneity. Similarly, many prominent rational asset-pricing models have been characterized as sensitive to "dark matter" that is difficult to measure with statistical accuracy (Chen, Dou and Kogan, 2024a).

This paper, instead, takes a "model-free" approach similar to Hansen and Jagannathan (1991) and Alvarez and Jermann (2005) but in a context where subjective beliefs may differ from rational expectations. Rather than specifying a particular model, I characterize some properties any asset-pricing model must have and in the process provide general estimates for the contributions of beliefs versus risks.

More specifically, I estimate the fraction of stochastic discount factor (SDF) volatility that originates from risks versus beliefs. Since SDFs summarize *all* asset prices, such model-free bounds and estimates should be useful to quantitative researchers seeking to align their models with data. Furthermore, the analysis provides a clear way to map financial survey data to key aspects of SDFs and other objects. Finally, approaching the problem at some level of generality can help unify disparate results from various specific markets. That is, rather than ask how much beliefs matter for, say, stock market volatility, this paper asks how much beliefs matter for asset pricing in general.

The basic identification problem. To account for large expected excess returns, we know that the SDF must be highly volatile (Hansen and Jagannathan, 1991). But what is not known is whether this volatility stems primarily from beliefs or risk. Ultimately, this identification problem is related to the fact that asset prices only encode the product of beliefs and risks and do not identify them separately (Harrison and Kreps, 1979). An Arrow-Debreu security for state x will have price

$$\text{Price}(x) = \text{Probability}(x) \times \underbrace{\text{Beliefs}(x) \times \text{MU}(x)}_{\text{SDF}(x)},$$

where $\text{Probability}(x)$ is the objective state probability, $\text{Beliefs}(x)$ denotes investors' belief distortion (ratio of subjective-to-objective probability), and $\text{MU}(x)$ is the representative investor's marginal utility in state x. It is this latter piece, $\text{MU}(x)$, that captures risk and risk attitudes. With enough data, perhaps the econometrician can figure out the state frequency $\text{Probability}(x)$. But beyond that, asset prices only identify the product $\text{Beliefs}(x) \times \text{MU}(x)$. To attribute SDF volatility to either component, asset prices do not suffice, and that is where financial surveys can help.

The literature now agrees further that most SDF volatility stems from permanent shocks (Alvarez and Jermann, 2005), i.e., from the first term in a decomposition like

$$\text{SDF} = \text{Permanent} \times \text{Transitory}. \tag{1}$$

Does this additional fact help us separate $\text{Beliefs}(x)$ from $\text{MU}(x)$? Unfortunately not. The *entirety* of $\text{Beliefs}(x)$ enters into the permanent component, so it cannot be distinguished from "long-term risks" to $\text{MU}(x)$ (Borovička, Hansen and Scheinkman, 2016). I show how, armed with survey data, one can partly separate long-term risks from beliefs.

Solution and results. It is relatively obvious that surveys provide *some information* about beliefs, but precisely what information and how should it be used? The core idea here, different from the literature, is that survey data on expected returns are connected to risk by no-arbitrage. If investors are optimistic about an asset, they must perceive it to be risky in some way. The result may seem counterintuitive: survey expected returns thus reveal something about investor marginal utility (MU), which captures risk sensitivities and is *distinct from beliefs*. On the other hand, realized returns contain information about the volatility of the SDF (Hansen and Jagannathan, 1991), which is affected by both risks and beliefs. By comparing survey expected returns to objective expected returns, one can thus learn how much of SDF volatility comes from MU volatility, i.e., how much stems from risk. I formalize this idea and explore variations in the paper.

More specifically, I make some simple, relatively model-free, assumptions and obtain various bounds of the following form:

$$\frac{\text{Volatility of MU}}{\text{Volatility of SDF}} \geq \frac{\text{Subjective risk premium}}{\text{Objective risk premium}}. \tag{2}$$

Take two extreme examples to build intuition. If investors held rational expectations, survey and objective expected returns would coincide. In that case, (2) implies that over 100% of SDF volatility is risk-based. Of course, this is the right answer: rational expectations leaves no room for belief distortions, and so the SDF is entirely risk-based. By contrast, if investors were risk-neutral, survey expected returns would equal risk-free rates. The right-hand-side of (2) would be 0%, consistent with the reality that all SDF volatility is due to beliefs, and none due to MU, under risk-neutrality.

My main empirical evidence, building on the logic above, suggests that at least half of total SDF volatility stems from risk. The key reason is the simple observation that subjective equity premia are large, almost as large as their true unconditional counterparts. In some variations, meaning either slight variations on the bound (2) or slight variations on the asset returns used, I find that over 100% of SDF volatility must be risk-based. Relatedly, Adam, Matveev and Nagel (2021) show that survey-respondents do not report risk-neutral expectations (meaning survey expected returns are significantly above riskless rates). I take their idea a step further and use (2) to *quantify* the fraction of SDF volatility due to risk.

What type of risk matters? I provide some direction by estimating a subjective version of the Alvarez and Jermann (2005) bounds. In particular, a decomposition like (1) exists for investors' MU as well as the SDF. I estimate that the permanent component of MU is responsible for at least 90% of total MU volatility. This cannot be attributed to beliefs, at least not directly, and it says that investors perceive substantial long-term risks.

Models and interpretation. Finally, I analyze several example models to illustrate and interpret the results. I consider a long-run risks model (Bansal and Yaron, 2004) and a growth extrapolation model with "perceived long-run risks" (Collin-Dufresne, Johannes and Lochstoer, 2017; Nagel and Xu, 2022; Bidder and Dew-Becker, 2016; Bhandari, Borovička and Ho, 2025). The latter is interesting because both beliefs and risks are operational: investors fear a hypothetical long-run risk, even though they are wrong. Although SDF volatility stems from an interaction between beliefs and preferences, I provide a clear classification of "belief-driven" versus "risk-driven" volatility. The calibrated growth extrapolation model is consistent with my estimated volatility bounds, revealing as a by-product that belief distortions can account for about 25% of SDF volatility. I also investigate what happens when the direct role for beliefs is magnified, by introducing sentiment shocks. In that case, the transitory component of the SDF is given too large a role, standing against existing evidence. Thus, within this class of models, the direct role of beliefs cannot be too large. Instead, if beliefs matter, they must matter mostly by creating marginal utility volatility, i.e., *subjective risks*. Isolating this more nuanced role for beliefs provides a potential way forward for the behavioral finance literature.

I conclude with an alternative model that features heterogeneous beliefs where investors "agree to disagree" (the closest counterpart is Dumas, Kurshev and Uppal, 2009). This exercise serves two purposes. First, I demonstrate how to aggregate beliefs and marginal utilities in a way that is both consistent with the "consensus belief" in surveys as well as my own volatility bounds. This allows me to reconcile, in a calibrated model, the large amount of disagreement in surveys with the minor role for belief distortions in SDF volatility. Second, the disagreement model reinforces my earlier conclusion about *subjective risks*. The SDF cannot inherit too much volatility directly from belief distortions, because that magnifies transitory SDF fluctuations. Instead, if beliefs matter significantly for the SDF, they must matter through induced marginal utility volatility.

Related literature and other approaches. My approach relies on volatility bounds, similar to Alvarez and Jermann (2005) and Bakshi and Chabi-Yo (2012). While this approach is relatively model-free, it can only provide partial information about SDFs, risks, and beliefs. If one wants more information, one needs more assumptions. For instance, assuming rational expectations and putting structure on objective dynamics allows the econometrician to extract investor marginal utility. An alternative puts some structure on marginal utility and treats beliefs as a "residual" needed to match asset market data (Ghosh and Roussellet, 2023; Chen, Hansen and Hansen, 2024c).

While these approaches are useful, my goal is to see how much can be said without these additional assumptions. I do not take a stand on investor preferences, the underlying objective dynamics, or the types of risks. Is there long-run growth risk; is uncertainty time-varying; does the economic state "jump" or evolve continuously? As long as we trust the survey data to proxy for investor beliefs, my procedure imposes minimal additional assumptions. The cost is that I can obtain only bounds on the volatilities of SDFs and marginal utilities, rather than their actual values.

A separate stream of literature is concerned with how belief dynamics may help resolve "excess volatility" puzzles, by combining present-value relations with survey data (Chen et al., 2013; De La O and Myers, 2021, 2024; De La O et al., 2023; Bordalo et al., 2019, 2024a,b; Bretscher et al., 2024; Cassella et al., 2025). This paper differs by focusing more on excess returns rather than excess volatility, and examining the SDF rather than a particular asset's present-value relation.[1] However, there is a connection between my work and this literature. If the market return is an important factor for the SDF, then its volatility is a key contributor to SDF volatility. In that case, subjective beliefs' role in SDF volatility and stock market volatility are linked, further reinforcing my conclusion that belief distortions must matter primarily via induced risk. Bretscher,

[1]While a decomposition into "future cash flows" and "future returns" is interesting in its own right, it is the wrong framework to evaluate whether return premia are due to behavioral biases or risk. For this question, it is better to work directly with returns and return expectations. For instance, it is not even necessary for valuation ratios or expected returns to be time-varying; even if those objects are constant, one can ask what fraction of constant premia arise from beliefs versus risk.

Malkhozov, Tamoni and Yang (2024) make a related point, that belief distortions show up in asset-price data largely as a risk factor.

My paper is mostly orthogonal to the behavioral excess volatility literature. For instance, many of my points are obtained simply from unconditional means, allowing me to sidestep a debate about the whether subjective expected returns are pro- or countercyclical (De Bondt, 1993; Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014; Nagel and Xu, 2023; Dahlquist and Ibert, 2024; Couts et al., 2024). Similarly, my example models feature little time-variation in valuation ratios, thus no excess volatility. And yet, those models allow beliefs to matter a lot for risk premia.

2 Theory: Subjective belief bounds

2.1 General environment

Let $\mathbb{P}$ denote the objective probability measure and $\mathbb{E}$ the associated expectation operator. I allow the marginal investor to possess *subjective beliefs* and use a distorted probability and expectation, which I denote by $\tilde{\mathbb{P}}$ and $\tilde{\mathbb{E}}$. To maintain a minimal amount of consistency for beliefs, I assume that $\mathbb{P}$ and $\tilde{\mathbb{P}}$ agree on null sets, so that their discrepancy can be modeled via the positive martingale B_t (likelihood ratio):

$$B_t = \left(\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right)_t \tag{3}$$

The subjective probability is defined by $\tilde{\mathbb{P}}(\mathcal{E}) := \mathbb{E}[\frac{B_t}{B_0}\mathbf{1}_{\mathcal{E}}]$ for any event $\mathcal{E}$ that resides in the time-t information set. This assumption on subjective beliefs is quite general and captures almost all belief-based models in the literature.[2]

I assume absence of arbitrage opportunities. (Because $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent measures, there is no ambiguity to simply saying "arbitrage" without reference to any probability.) In that case, let $(S_t)_{t\geq 0}$ denote the *stochastic discount factor* (SDF) process that prices all assets, i.e., S is positive and for any gross return R_{t+1} we have

$$1 = \mathbb{E}_t\Big[\frac{S_{t+1}}{S_t}R_{t+1}\Big], \tag{4}$$

[2]Rule out are models where the law of iterated expectations (LIE) is violated (e.g., Fuster et al., 2010). Other models, like the fading memory model of Nagel and Xu (2022), may violate LIE when interpreted literally but can be written in an observationally equivalent form where LIE holds. Therefore, that model would be included in the general setup here.

where $\mathbb{E}_t$ denotes the conditional expectation operator. At the same time, there is a subjective SDF $\tilde{S}_t$ that prices assets under investor beliefs, i.e., investors' Euler equation:

$$1 = \tilde{\mathbb{E}}_t\Big[\frac{\tilde{S}_{t+1}}{\tilde{S}_t}R_{t+1}\Big]. \tag{5}$$

The interpretation of equation (5) is that $\tilde{S}_t$ represents the *investors' marginal utility of wealth*—I refer to $\tilde{S}$ as marginal utility, in contrast to the objective SDF S. There are two key economic assumptions embedded in (5). First, I am assuming that some investors are marginal in all the relevant markets considered in this paper. Second, because I identify the survey results with the subjective probability $\tilde{\mathbb{P}}$, I am essentially assuming that these marginal investors hold beliefs identical to those in the surveys.

Without loss of any generality for my points, impose the normalization $S_0 = \tilde{S}_0 = B_0 = 1$. Then, comparing (4)-(5) and using (3), we have

$$S_t = \tilde{S}_t B_t. \tag{6}$$

The SDF is the product of investor marginal utility and investor beliefs. Roughly speaking, the idea of this paper is to use survey data to learn about the volatilities of S and $\tilde{S}$ separately, thereby providing information about the belief distortion B.

In addition to the volatilities of S and $\tilde{S}$, information is contained in the volatilities of their permanent components. Under fairly general conditions, which I assume, the SDF obeys a unique factorization into *permanent and transitory components* (Alvarez and Jermann, 2005; Hansen and Scheinkman, 2009; Qin and Linetsky, 2016, 2017):

$$S_t = G_t H_t, \tag{7}$$

where G is trend-stationary (transitory component) and H is a $\mathbb{P}$-martingale (permanent component). An analogous factorization holds for investors' marginal utility:

$$\tilde{S}_t = \tilde{G}_t \tilde{H}_t, \tag{8}$$

where $\tilde{G}$ is trend-stationary and $\tilde{H}$ is a $\tilde{\mathbb{P}}$-martingale.

Luckily, equation (6) then implies that

$$H_t = \tilde{H}_t B_t \tag{9}$$

and hence $G_t = \tilde{G}_t$. To see this, notice that $\mathbb{E}_0[H_t] = H_0\mathbb{E}_0[\frac{B_t}{B_0}\frac{\tilde{H}_t}{\tilde{H}_0}] = H_0\tilde{\mathbb{E}}_0[\frac{\tilde{H}_t}{\tilde{H}_0}] = H_0$, so that H and $\tilde{H}$ are indeed $\mathbb{P}$- and $\tilde{\mathbb{P}}$-martingales, respectively, when they satisfy (9). The interpretations of $\tilde{H}$ and H are different: $\tilde{H}$ captures the impact of permanent risks on investor marginal utility, while $H = \tilde{H}B$ captures the joint impact of permanent risks and belief distortions. The permanent components H and $\tilde{H}$ deserve attention, in light of evidence from Alvarez and Jermann (2005) and Bakshi and Chabi-Yo (2012) that G carries a relatively minor contribution to overall SDF volatility. What is not yet agreed is whether the prominence of H is due to $\tilde{H}$ or B.

This entire theoretical setup has referenced a single investor belief and marginal utility, when in fact there is clearly heterogeneity in the real world and in the surveys that I use empirically. Appendix D shows that the presence of heterogeneity can be accommodated in this setup. In particular, the common practice of averaging across survey respondents (the so-called consensus forecast) allows us to interpret B, $\tilde{S}$, and $\tilde{H}$ as the average belief, the belief-weighted average marginal utility, and its permanent component. The example structural models in Section 5 feature disagreement and allow me to provide more interpretation for these weighted-average aggregates.

Finally, I will make frequent reference to a few special returns in this theory section. First, I denote the *one-period risk-free rate* R_t^f. Since this is conditionally risk-free, we have from (4)-(5) that $R_t^f = (\mathbb{E}_t \frac{S_{t+1}}{S_t})^{-1} = (\tilde{\mathbb{E}}_t \frac{\tilde{S}_{t+1}}{\tilde{S}_t})^{-1}$. I will assume, as in most of the asset-pricing literature, that the risk-free rate is traded.

Second, I will make use of the *long bond return* R_{t+1}^∞, defined as the one-period holding return on an arbitrarily long-maturity discount bond:

$$R_{t+1}^\infty := \lim_{T\to\infty} \frac{\mathbb{E}_{t+1}[\frac{S_T}{S_{t+1}}]}{\mathbb{E}_t[\frac{S_T}{S_t}]}.$$

The long bond was originally studied in Kazemi (1992) (see also Martin and Ross, 2013, and Ross, 2015). In particular, it turns out that R_{t+1}^∞ identifies the transitory component of the SDF:

$$R_{t+1}^\infty = \lim_{T\to\infty} \frac{\mathbb{E}_{t+1}[\frac{S_T}{S_{t+1}}]}{\mathbb{E}_t[\frac{S_T}{S_t}]} = \lim_{T\to\infty} \frac{G_t}{G_{t+1}} \frac{\mathbb{E}_{t+1}[\frac{H_T}{H_{t+1}} G_T]}{\mathbb{E}_t[\frac{H_T}{H_t} G_T]} = \frac{G_t}{G_{t+1}}, \tag{10}$$

because the ratio of conditional expectations converges to 1 under mild conditions. Of course, an infinite-maturity discount bond does not exist, but we can proxy this with the longest-maturity bonds available.

Third, I will occasionally consider the *growth-optimal return* R_{t+1}^*. This is the portfolio that maximizes the growth rate of wealth, i.e., it solves $\max_R \mathbb{E}_t[\log(R)]$ subject to the pricing constraint (4). The solution is $R^*_{t+1} = \frac{S_t}{S_{t+1}}$ (Roll, 1973; Bansal and Lehmann, 1997). Thus, there is a well-known equivalence between extracting the SDF from the data and finding the growth-optimal portfolio. Because of the difficulties in identifying R^*, I will not rely on any one particular approach but rather present results from several. Note that there is also a *perceived growth-optimal return* $\tilde{R}^*_{t+1}$, which the subjective version and is defined analogously by $\arg\max_R \tilde{\mathbb{E}}_t[\log(R)]$. I will also refer to $\tilde{R}^*_{t+1}$ at times.

2.2 Volatility bounds

I consider *entropy measures of volatility*. For any positive random variable X, define

$$L_t(X) := \log\left(\mathbb{E}_t[X]\right) - \mathbb{E}_t\left[\log(X)\right] \tag{11}$$

$$\tilde{L}_t(X) := \log\left(\tilde{\mathbb{E}}_t[X]\right) - \tilde{\mathbb{E}}_t\left[\log(X)\right]. \tag{12}$$

Entropy measures played an important role in Alvarez and Jermann (2005), but have also been used in various other asset-pricing contexts (Backus et al., 2011, 2014; Martin, 2017) By Jensen's inequality, $L_t(X) \geq 0$ and $\tilde{L}_t(X) \geq 0$. In the special case that X_{t+1} is conditionally lognormal, $L_t(X_{t+1}) = \frac{1}{2}\text{Var}_t[\log X_{t+1}]$, but in other cases entropy conveys more information about the left tail of a distribution, relative to variance.

Whereas $L_t(X)$ represents objective volatility, $\tilde{L}_t(X)$ represents subjective, or perceived, volatility. To make interpretable comparisons, I will sometimes invoke the following condition relating objective and subjective entropies. This condition allows us to bound time-series averages of conditional entropies.

Definition 1. X satisfies the *unconditional entropy inequality* (UEI) if $L(X_{t+1}) \geq \mathbb{E}[\tilde{L}_t(X_{t+1})]$.

When does a variable X to satisfy UEI? To understand this condition, first note the following property, analogous to the "law of total variance" but for entropies:

$$L(X) = \mathbb{E}[L_t(X)] + L(\mathbb{E}_t[X])$$

Total entropy is the average of conditional entropy plus the entropy of the conditional mean. With this law in mind, and noting that $L(\mathbb{E}_t[X]) \geq 0$, one obvious situation where UEI holds is if $L_t(X) \geq \tilde{L}_t(X)$, meaning that the subjective measure understates risk to X. This case is relevant because, empirically, survey measures of risk tend to be lower than actual risk (see Figure 1). This case is also theoretically appealing, because when the underlying joint dynamics are conditionally lognormal over sufficiently short time periods (i.e., no jumps), then $L_t(X) = \tilde{L}_t(X)$ and so UEI holds.[3] Most models in macro-finance have this property. A second situation where UEI is likely appropriate is if $L(\mathbb{E}_t[X])$ is large, meaning that X has significant predictability. This case is relevant because actual returns do tend to have some predictability.

Lemma 1. *The following conditional volatility bounds hold for all returns R_{t+1}:*

$$L_t\left(\frac{S_{t+1}}{S_t}\right) \geq \mathbb{E}_t\left[\log R_{t+1}\right] - \log R_t^f \tag{13}$$

$$\tilde{L}_t\left(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}\right) \geq \tilde{\mathbb{E}}_t\left[\log R_{t+1}\right] - \log R_t^f \tag{14}$$

and

$$L_t\left(\frac{H_{t+1}}{H_t}\right) \geq \mathbb{E}_t\left[\log R_{t+1} - \log R_{t+1}^{\infty}\right] \tag{15}$$

$$\tilde{L}_t\left(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}\right) \geq \tilde{\mathbb{E}}_t\left[\log R_{t+1} - \log R_{t+1}^{\infty}\right] \tag{16}$$

The proofs of all theoretical results are contained in Appendix A.

The SDF bound (13) is exactly the "growth-optimal bound" uncovered by Bansal and Lehmann (1997). Expected log returns, which incorporate a risk adjustment due to the concavity of log, bound the entropy of the SDF from below. Viewed this way, the SDF bound is also reminiscent of how risk-adjusted returns bound SDF volatility in Hansen and Jagannathan (1991). What is new is the marginal utility bound (14). Whereas the first object $L_t(\Delta S_{t+1})$ measures the true conditional volatility of $S = B\tilde{S}$, which jointly captures contributions from beliefs (B) and risks ($\tilde{S}$), the second object $\tilde{L}_t(\Delta \tilde{S}_{t+1})$ isolates the role of risks. If conditional log-normality holds and time-periods are short, then $\tilde{L}_t(\Delta \tilde{S}_{t+1}) \approx L_t(\Delta \tilde{S}_{t+1})$, so that the perceived marginal utility volatility equals true marginal utility volatility. In that case, comparing the volatilities from (13)-(14) provides information about the importance of the belief distortion B, which is the wedge between S and $\tilde{S}$. This is the basic idea behind the entire paper.

A similar discussion holds for the permanent components H and $\tilde{H}$. Whereas (15) measures the total volatility coming from beliefs (B) and long-run risks ($\tilde{H}$) together, the subjective bound (16) measures the perceived conditional volatility of the long-run risks ($\tilde{H}$) alone. The gap between them informs us about how important beliefs alone are to asset pricing.

[3] One easy way to see this is that $L_t(X_{t+1}) = \frac{1}{2}\text{Var}_t[\log X_{t+1}]$ and $\tilde{L}_t(X_{t+1}) = \frac{1}{2}\tilde{\text{Var}}_t[\log X_{t+1}]$ under conditional lognormality. Then, combine this result with the fact that belief distortions about variance necessarily become negligible as the time interval shrinks. See Appendix A.6 for a formal proof.

At first glance, it is not obvious why $\tilde{\mathbb{E}}_t[\log R_{t+1}] - \log R_t^f$ and $\tilde{\mathbb{E}}_t[\log R_{t+1} - \log R_t^\infty]$ are informative about risks, rather than beliefs. Is it not possible that investors are just optimistic about an asset return R, which has nothing to do with their marginal utility $\tilde{S}$? No-arbitrage rules this out: if investors are optimistic about an asset, they must perceive it to have *some* risk. Since marginal utility $\tilde{S}$ summarizes investors' risk attitudes, this risk necessarily manifests in their marginal utility volatility.

In light of this discussion, let us pause to build intuition by reconsidering a question asked in Adam et al. (2021): do survey-respondents report risk-neutral expectations, and if so what does that imply? If $\tilde{\mathbb{E}}_t[R_{t+1}] = R_t^f$, then $\tilde{\mathbb{E}}_t[\log R_{t+1}] - \log R_t^f \leq 0$. If this holds for all assets R, then the lower bound in (14) would be tight for $R_{t+1} = R_t^f$, meaning that $\tilde{L}_t(\Delta \tilde{S}_{t+1}) = 0$. In this hypothetical world, we would conclude that *none* of SDF volatility comes from risk, and *all* SDF volatility emanates from beliefs. Essentially, this is just a restatement of the fact that, if they survey-takers are marginal investors, then we would conclude from the surveys that investors are risk-neutral, i.e., $\Delta \tilde{S}_{t+1} = 1/R_t^f$. Of course, Adam et al. (2021) reject this hypothesis empirically, meaning that *at least some* SDF volatility must come from risk. One can view this paper as taking their idea one step further and quantifying *how much* of SDF volatility comes from risk.

One additional benefit of working under subjective beliefs is the direct availability of conditional expectations. In particular, survey data on subjective expected returns contain proxies for $\tilde{\mathbb{E}}_t[\log R_{t+1}]$ and $\tilde{\mathbb{E}}_t[\log R_{t+1}^\infty]$, which combined with Lemma 1 delivers time series of bounds for $\tilde{L}_t(\Delta \tilde{S}_{t+1})$ and $\tilde{L}_t(\Delta \tilde{H}_{t+1})$. By contrast, it is often more difficult to find a reasonable proxy for objective conditional expectations of returns.

Building on the conditional results of Lemma 1, I also present unconditional bounds. For the objective probability, this is straightforward. On the other hand, the subjective bound does not generalize immediately to a useful unconditional version. Indeed, the *unconditional* subjective expectation $\tilde{\mathbb{E}}$ is not directly observable; taking time-series averages of *conditional* subjective expectations $\tilde{\mathbb{E}}_t$ will not yield an estimate of $\tilde{\mathbb{E}}$ unless beliefs are rational. We can make progress if we invoke UEI in Lemma 1.

Proposition 1. *The following unconditional volatility bounds hold for all returns R_{t+1}:*

$$L\Big(\frac{S_{t+1}}{S_t} R_t^f\Big) \geq \mathbb{E}\big[\log R_{t+1} - \log R_t^f\big] \tag{17}$$

$$L\Big(\frac{H_{t+1}}{H_t}\Big) \geq \mathbb{E}\big[\log R_{t+1} - \log R_{t+1}^\infty\big] \tag{18}$$

Additionally, if UEI holds for $\frac{\tilde{S}_{t+1}}{\tilde{S}_t}$ and $\frac{\tilde{H}_{t+1}}{\tilde{H}_t}$, then

$$L\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}R_t^f\Big) \geq \mathbb{E}\Big[\tilde{\mathbb{E}}_t\big[\log R_{t+1} - \log R_t^f\big]\Big] \tag{19}$$

$$L\Big(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}\Big) \geq \mathbb{E}\Big[\tilde{\mathbb{E}}_t\big[\log R_{t+1} - \log R_{t+1}^\infty\big]\Big] \tag{20}$$

Finally, while Lemma 1 and Proposition 1 hold in levels, we may want a ratio version in order to lower-bound the fraction of long-run SDF volatility that is due to marginal utility rather than beliefs. This approach requires more assumptions—namely the hypothesis that a return is "closer" to the objective growth-optimal return than a subjective counterpart (which holds, for example, if the chosen return is actually the growth-optimal return, but in other cases too). But the benefit is this approach permits a more intuitive scale as a ratio of volatilities.

Definition 2. For any return R, define its *distance to growth-optimal* $\delta_t(R)$ and *distance to perceived growth-optimal* $\tilde{\delta}_t(R)$, respectively, by

$$\delta_t(R) := \mathbb{E}_t[\log R_{t+1}^*] - \mathbb{E}_t[\log R_{t+1}]$$
$$\tilde{\delta}_t(R) := \tilde{\mathbb{E}}_t[\log \tilde{R}_{t+1}^*] - \tilde{\mathbb{E}}_t[\log R_{t+1}].$$

Proposition 2. *Suppose UEI holds for $\frac{\tilde{S}_{t+1}}{\tilde{S}_t}$ and $\frac{\tilde{H}_{t+1}}{\tilde{H}_t}$. Let R and $\tilde{R}$ be any two returns such that, on average, R is closer to growth-optimal than $\tilde{R}$ is to perceived growth-optimal, in the sense that $\mathbb{E}[\delta_t(R)] \leq \mathbb{E}[\tilde{\delta}_t(\tilde{R})]$. Then, the following volatility bounds hold:*

$$\frac{L(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}R_t^f)}{L(\frac{S_{t+1}}{S_t}R_t^f)} \geq \min\left\{1,\ \frac{\mathbb{E}\big[\tilde{\mathbb{E}}_t[\log \tilde{R}_{t+1} - \log R_t^f]\big]}{\mathbb{E}[\log R_{t+1} - \log R_t^f]}\right\} \tag{21}$$

$$\frac{L(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{L(\frac{H_{t+1}}{H_t})} \geq \min\left\{1,\ \frac{\mathbb{E}\big[\tilde{\mathbb{E}}_t[\log \tilde{R}_{t+1} - \log R_{t+1}^\infty]\big]}{\mathbb{E}[\log R_{t+1} - \log R_{t+1}^\infty]}\right\} \tag{22}$$

The results so far allow me to infer how much risk matters for the SDF, but what type of risk matters? I conclude this section with two auxiliary results, stated as corollaries, that measure the size of the permanent components of the SDF and marginal utility—the first one due to Alvarez and Jermann (2005), and the second one a novel subjective (and conditional) version. Whereas the Alvarez and Jermann (2005) bound quantifies the permanent component H, the subjective version below quantifies the subjective permanent component $\tilde{H}$, which is informative about how much risk investors perceive as coming from permanent sources. I will also report empirical estimates for these bounds.

Corollary 1.

$$\frac{L\left(\frac{H_{t+1}}{H_t}\right)}{L\left(\frac{S_{t+1}}{S_t}\right)} \geq \min\left\{1, \frac{\mathbb{E}\left[\log R_{t+1} - \log R^{\infty}_{t+1}\right]}{\mathbb{E}\left[\log R_{t+1} - \log R^f_t\right] + L(1/R^f_t)}\right\} \tag{23}$$

for any return R such that $\mathbb{E}[\log R_{t+1} - \log R^f_t] + L(1/R^f_t) > 0$.

Corollary 2.

$$\frac{\tilde{L}_t\left(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}\right)}{\tilde{L}_t\left(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}\right)} \geq \min\left\{1, \max\left\{0, \frac{\tilde{\mathbb{E}}_t\left[\log R_{t+1} - \log R^{\infty}_{t+1}\right]}{\tilde{\mathbb{E}}_t\left[\log R_{t+1}\right] - \log R^f_t}\right\}\right\} \tag{24}$$

for any return R such that $\tilde{\mathbb{E}}_t \log R_{t+1} > \log R^f_t$.

3 Estimating the bounds

The results to follow use the following return notations. Let R^m denote the aggregate stock market return. Given a positive number k, let y^k and R^k denote the log yield and holding return, respectively, on a discount bond with maturity k. For example, R^{10} denotes the return on a 10-year Treasury. The risk-free rate between t and $t+1$ will alternatively be written either as R^1_{t+1}, following the notation for other bonds, or as R^f_t, following the convention from the theory section. I assume a period length of one year, so the risk-free rate corresponds to the yield of a 1-year Treasury security.

3.1 Data and variable construction

Table F.1 in Appendix F.1 summarizes the various survey and asset market data sources.

For the realized returns and yields, the data sources are standard. Realized returns on the aggregate value-weighted stock market are obtained from CRSP. I supplement this with a few volatility measures: I construct realized variance from daily return series obtained from Ken French's website; I use the CBOE VIX obtained from OptionMetrics; and I use the "SVIX" from Ian Martin's website (Martin, 2017). Realized Treasury yields and returns come from the St. Louis Fed (FRED), the Fed Board updates to Gürkaynak et al. (2007) (GSW), and CRSP. When there is overlapping data on par yields, I use the FRED yields as the first choice, followed by the GSW second, followed by the CRSP last.

For zero-coupon yields, GSW is the only source. Appendix F.2 displays these various sources for comparison. For bond returns, CRSP is the only source. From the fixed-term indexes file, I obtain return series corresponding to an approximately 10-year and 30-year bond, respectively. I also obtain returns on a bond index tracking bonds with maturities above 10 years.

Survey-based US stock market expected returns come from Nagel and Xu (2023), who compile three sources: (i) the individual investor expectations from Nagel and Xu (2022), which agglomerate the UBS/Gallup survey, the Conference Board survey, and the Michigan Survey of Consumers, plus a few smaller surveys; (ii) the Livingston Survey from the Philadelphia Fed; and (iii) the Graham-Harvey CFO survey (Ben-David et al., 2013). From these sources, Nagel and Xu (2023) construct proxies for annual expected excess returns on the stock market (or S&P 500, depending on the source). See Nagel and Xu (2023) for more details on this data.

I convert expected stock returns $\tilde{\mathbb{E}}_t[R^m_{t+1}]$ to expected log returns $\tilde{\mathbb{E}}_t[\log R^m_{t+1}]$ via the identity $\tilde{\mathbb{E}}_t[\log R_{t+1}] = \log \tilde{\mathbb{E}}_t[R_{t+1}] - \tilde{L}_t(R_{t+1})$. While I do not have access to a measure of perceived entropy $\tilde{L}_t(R_{t+1})$, the lognormal approximation implies $\tilde{L}_t(R_{t+1}) \approx \frac{1}{2}\tilde{\mathrm{Var}}_t[\log R_{t+1}]$. Nagel and Xu (2023) construct a measure of $\tilde{\mathrm{Var}}_t[\log R^m_{t+1}]$ from the elicited 10th and 90th percentile return forecasts in the CFO survey. To extend this subjective variance measure to a longer history, I regress it on measures of uncertainty that are available further back in time. First, I regress the subjective variance $\tilde{\mathrm{Var}}_t[\log R^m_{t+1}]$ onto the contemporaneous squared VIX, its one-month lag, and its trailing-12-month moving average. This regression has an R-squared of 0.60 and generates the fitted value $\hat{y}_t = 0.012 + 0.52\mathrm{VIX}^2_t + 0.005\mathrm{VIX}^2_{t-1} + 0.071(\frac{1}{12}\sum_{j=0}^{11}\mathrm{VIX}^2_{t-j})$.[4] I then backfill fitted values for months that pre-date the VIX by using the Nagel and Xu (2023) procedure of projecting VIX onto the news-implied volatility (NVIX) measure of Manela and Moreira (2017). The result of this is a time series of fitted and backfilled subjective variance estimates $\tilde{\mathrm{Var}}_t[\log R^m_{t+1}]$ that I use to construct expected log market returns as $\tilde{\mathbb{E}}_t[\log R^m_{t+1}] = \log \tilde{\mathbb{E}}_t[R^m_{t+1}] - \frac{1}{2}\tilde{\mathrm{Var}}_t[\log R^m_{t+1}]$. The results are displayed in Figures 1-2.

Figure 1 displays the CFO's reported subjective volatility as well as its fitted value from the projection onto VIX and lags. The first thing to notice is that the subjective

[4] The VIX is theoretically a reasonable regressor. Indeed, Result 3 of Martin (2017) shows that $\mathrm{VIX}^2_{t\to T} = \frac{2}{T-t}L^*_t(R^m_{t\to T})$, where L^* is the risk-neutral entropy. To the extent that the risk-neutral and investors' subjective entropies move together at the one-year horizon (as they would in a conditionally log-normal model), the VIX is the ideal regressor. That said, the standard VIX is a one-month-ahead implied volatility, while the independent variable $\tilde{\mathrm{Var}}_t[\log R^m_{t+1}]$ is a one-year ahead volatility. However, this seems to be of minor importance: adding the one-year squared SVIX (Martin, 2017) to this regression produces a very minor adjustment to the resulting fitted value of CFO perceived return volatility—in particular, for the sample where SVIX is available, I find an R-squared of 0.778 from including it, versus 0.774 without it.

volatility measure shares many of the same patterns as objective volatility measures—recall the R-squared of 0.60 mentioned above in the fitting procedure. That said the spikes in the subjective volatility are milder than the objective volatility spikes. On average, the subjective CFO volatility (13.2%) is substantially below the VIX (20.9%) and slightly below realized volatility (16.6%). Note that this is consistent with the UEI condition: recall that $\mathbb{E}[L_t(X)] \geq \mathbb{E}[\tilde{L}_t(X)]$ is a sufficient condition for UEI; taking variance as a good proxy for risk perception, the data suggest that this sufficient condition holds.

Figure 2 displays the ultimate resulting expected excess log stock returns after making the variance correction described above. There are some shared cyclical patterns but also discrepancies. Importantly, my results primarily make use of the average levels of these log returns rather than their volatilities or cyclical dynamics (except insofar as those second moments impact standard errors).

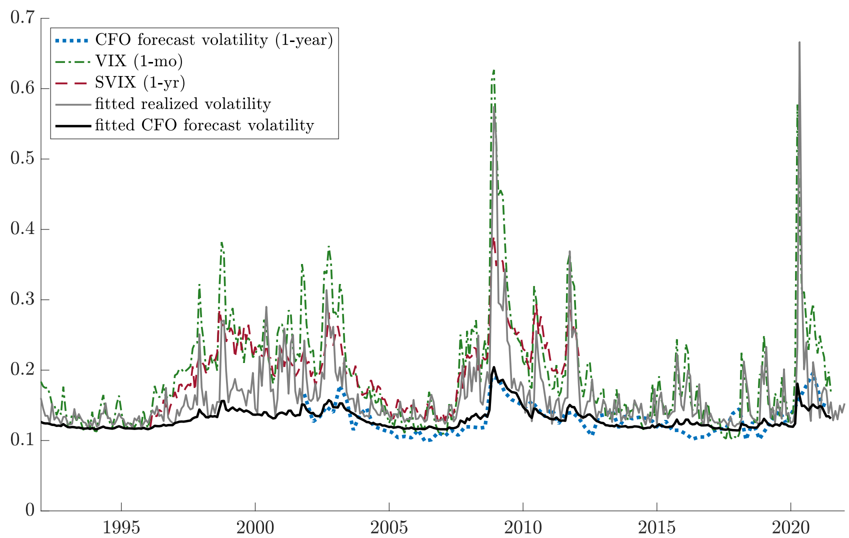

Figure 1. Subjective and Objective Volatilities. This figure plots fitted CFO volatility (solid black) from a regression of CFO's subjective variance forecasts on the squared VIX and lags of squared VIX. The "fitted realized volatility" is the square root of the predicted value from an AR(1) model for realized variance of the value-weighted market portfolio.

Treasury Bond expected returns are constructed from yield forecasts in the Blue Chip Financial Forecasts (BCFF) survey. BCFF has consensus (i.e., averaged across respondents) yield forecasts for Treasuries with maturities 6 months, 1 year, 2 years, 5 years, 10 years, and 30 years.[5] I follow Kim and Orphanides (2012) in treating the BCFF forecasts as approximately mid-quarter forecasts, and so I use a weighted-average of the 3-, 4-, and 5-quarter ahead forecasts, with weights depending on the survey calendar month, to get

[5] From March 2002 through February 2006, there was no 30-year bond issued, so the survey records a forecast of the 20-year yield. I account for this in all results with both the survey expected returns and realized returns.

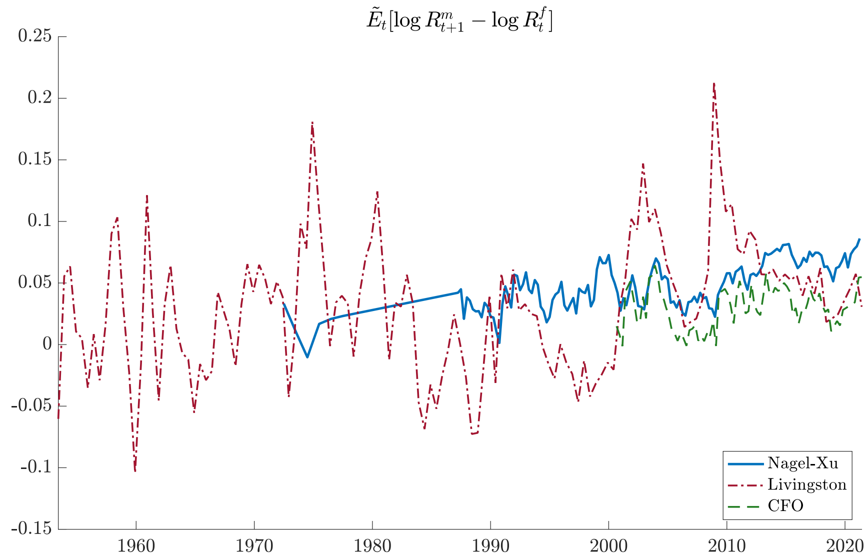

Figure 2. Subjective expected excess stock returns. This figure plots survey-based expected one-year log stock returns in excess of the one-year Treasury rate. Expected log returns are obtained from expected arithmetic returns via a variance adjustment that is described in the text.

a forecast at a one-year horizon. Since the BCFF surveys are taken nearer to the beginning of the month, I also shift the responses and align them with with previous-month yields and asset prices, which are measured at month-end (this shift does not materially affect any result). I then interpolate these yields to obtain a full term structure by linear interpolation for maturities up to 5 years and a fitted Nelson-Siegel model for maturities beyond 5 years (with the Nelson-Siegel model fitted to expectations data on maturities 2+ years). The result is a term structure of one-year-ahead expectations of the par yield curve. I then bootstrap the approximate expected continuously compounded zero-coupon yield curve $(\tilde{\mathbb{E}}_t[y^k_{t+1}])^{30}_{k=1}$ from this expected par yield curve.[6] Appendix F.3 contains additional documentation of the transformation from the raw data (sparse par yield forecasts) to an interpolated par yield curve forecast to a bootstrapped zero-coupon yield curve forecast. Finally, I construct one-year subjective expected log returns via $\tilde{\mathbb{E}}_t[\log R^k_{t+1}] = ky^k_t - (k-1)\tilde{\mathbb{E}}_t[y^{k-1}_{t+1}]$ at the $k = 10$-year and $k = 30$-year maturities.

As an extension, I also project the bond expectations back and forward in time by regressing them on various yields and their lags.[7] I regress $\tilde{\mathbb{E}}_t y^9_{t+1}$ (resp., $\tilde{\mathbb{E}}_t y^{29}_{t+1}$) onto the time-$t$ 1-year and 10-year (30-year) par yields, as well as their one-month lagged values and a trailing-12-month average of their values. The in-sample R-squareds from

[6]Technically, this bootstrapping is a nonlinear transformation. In order to construct the zero-coupon yield expectations using this approximation, I am effectively assuming the amount of uncertainty in the par yield expectations is small.

[7]I am in the process of obtaining a longer time series of BCFF data and will de-emphasize this backfilling in a subsequent version.

these regressions are 0.9821 and 0.9612, respectively. I use this estimation to construct hypothetical values of $\tilde{\mathbb{E}}_t[\log R^{10}_{t+1}]$ and $\tilde{\mathbb{E}}_t[\log R^{30}_{t+1}]$ going back as far as I have actual yield data.

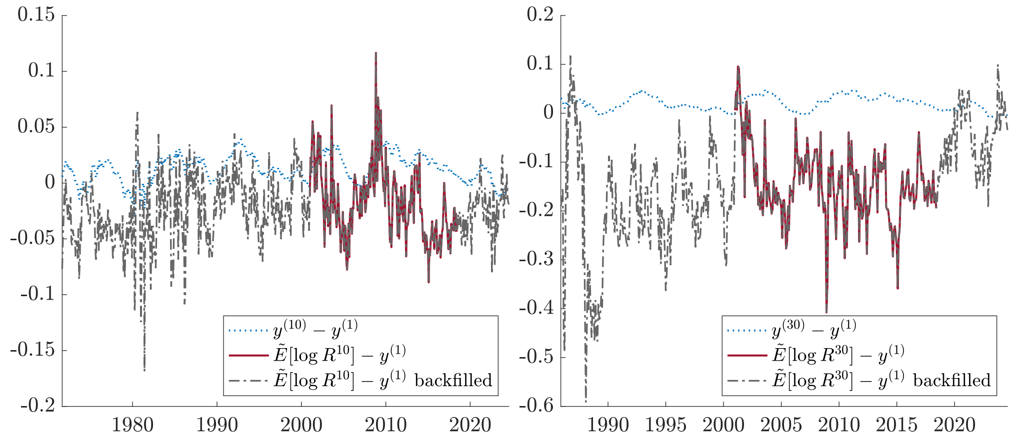

Figure 3. Subjective expected bond returns. Expected log (zero-coupon) bond returns constructed using BCFF consensus forecasts of the one-year ahead yield curve. "Backfilled" refers to expected returns constructed from projecting yield forecasts onto the contemporaneous and lagged yield curve.

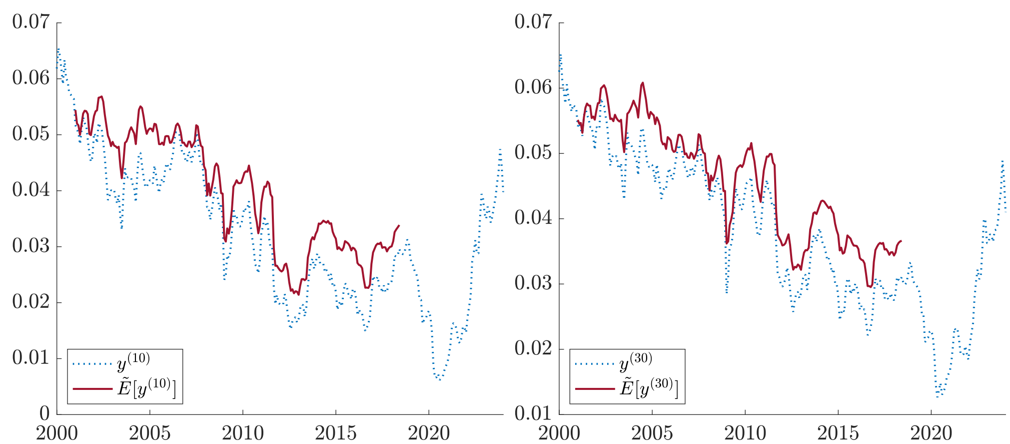

Figure 4. Yields and yield forecasts. Par yields and the BCFF consensus one-year ahead forecasts of those par yields. All yields and yield forecasts are converted from coupon-equivalent to continuously compounded units.

The resulting subjective expected excess log bond returns, as well as their extended backfilled values, are depicted in Figure 3. Notice from the figure that the subjective expected returns tend to be below the long-maturity yield. Since the yield measures the long-term expected return, this implies that survey respondents are pessimistic about bond returns in this sample. The situation is particularly extreme in 30-year bonds. As Figure 4 shows, this stems from pessimism about yields: survey respondents think yields will either stop falling or even rise during this period. The extremely pessimistic

30-year bond expected return is then the result of general pessimism about yields (Figure 4 shows that both the 10-year and 30-year yield forecasts are consistently about 0.5% above yield levels) combined with the significantly higher duration of the 30-year bond.

3.2 Permanent versus transitory shocks

To start, Table 1 reviews and refreshes the original Alvarez and Jermann (2005) results on the size of the permanent component in the SDF. Overall, the message from the point estimates is that at least 2/3 of the volatility of S stems from H. The reason: the equity premium (5.55%) is substantially larger than the term premium (between 1% and 2%, depending on the method), so that $\mathbb{E}[\log R^m - \log R^\infty]$ is likely to be nearly as large as the entire equity premium. Standard errors, shown in parentheses, are computed using a Newey and West (1987) window with 36 months of lags to account for the overlap in returns and persistence in spreads. The standard errors for the size of the permanent component are computed using the Delta Method. Based on these standard errors, the estimates appear to be reasonably precise.

As in the original paper, I use several methods for the long bond component. I use both holding period returns (Panel A) and yields (Panel B) as alternatives to measure expected returns. In a stationary and ergodic environment, the time-series average yield also measures a bond expected return; yields are much less volatile (though more persistent) than holding returns, which explains why Panel B features much tighter standard errors. I also examine both the 10-year and 30-year maturities as proxies for the long bond. While 30-year bonds are preferable in principle, the discrepancy in results is negligible in practice.

Next, I perform a similar analysis in the subjective measure. Table 2 presents estimates for the average size of the permanent component $\tilde{H}$ in marginal utility $\tilde{S}$. The estimates in the final column are the time-series average of the conditional subjective bound from Corollary 2. As before, standard errors, shown in parentheses, are computed using a Newey and West (1987) window with 3 years of lags; since the equity surveys feature different sampling frequencies, this means a different number of period lags in each panel (Nagel-Xu and CFO data use 12 quarter lags; the Livingston survey use 6 semi-annual lags). The results, across the three equity surveys and four different proxies for the long bond expected return, broadly echo the message of Alvarez and Jermann (2005): investors perceive that the lion's share of marginal utility volatility (90% or more) emanates from permanent risks.

In my sample, I find a very negative subjective long bond term premium, particularly

	Equity Premium $E[\log R^m/R^1]$	Term Premium $E[\log R^{(k)}/R^1]$	Adjustment for Volatility of Short Rate $L(1/R^1)$	Size of Permanent Component $L(\Delta H)/L(\Delta S)$
A. Holding returns.				
$k = 10$ years	0.0555	0.0100	0.0005	0.8134
	(0.0161)	(0.0099)	(0.0001)	(0.1641)
$k = 30$ years		0.0091		0.8292
		(0.0143)		(0.2308)
$k > 10$ (index)		0.0125		0.7686
		(0.0087)		(0.1564)
B. Yields.				
$k = 10$ years (par)	0.0555	0.0097	0.0005	0.8185
	(0.0161)	(0.0016)	(0.0001)	(0.0566)
$k = 10$ years (zcb)		0.0129		0.7615
		(0.0022)		(0.0718)
$k = 30$ years (par)		0.0119		0.7791
		(0.0027)		(0.0692)
$k = 30$ years (zcb)		0.0195		0.6428
		(0.0031)		(0.1096)

Table 1. Size of permanent shocks in the SDF. This table displays estimates of the objective volatility bound from Corollary 1, originally due to Alvarez and Jermann (2005). Panel A uses holding period returns to measure the term premium; each row uses a different proxy for R^∞. Panel B uses yields (par or zero-coupon bond yields) to measure the term premium. Newey-West asymptotic standard errors using 36 months of lags are shown in parentheses.

for the 30-year bond. This is the source of the large estimated subjective permanent component. This could be sample-specific: the BCFF sample period here (2001–2018) only covers a period of low long-term yields, which survey respondents expected to revert, leading to extremely low subjective term premia. The backfilled yield forecasts, which lead to very similar results in Table 2, may not address this issue because they inherit the estimated in-sample pessimism from the BCFF survey.[8]

At this point, we know that (i) permanent shocks are the dominant source of SDF variation, and (ii) permanent risks are the dominant source of marginal utility variation. These findings can be rationalized in one of two ways. It is possible that agents have nearly-rational expectations, meaning all the permanent shocks and SDF variation correspond closely to permanent risks and marginal utility variation, respectively. At

[8] For example, when projecting the 29-year zero-coupon yield forecast onto the 1-year and 30-year par yields, their one-month lags, and their trailing-12-month average, I estimate

$$\widehat{\tilde{\mathbb{E}}_t[y^{(29)}_{t+1}]} = -0.074 + 0.316Y^{(1)}_t + 0.614Y^{(30)}_t - 0.097Y^{(1)}_{t-1} + 0.182Y^{(30)}_{t-1} - 0.434\left(\frac{1}{12}\sum_{j=0}^{11} Y^{(1)}_{t-j}\right) + 0.518\left(\frac{1}{12}\sum_{j=0}^{11} Y^{(30)}_{t-j}\right)$$

with an R-squared of 0.9612. The negative estimated constant can be interpreted as estimated pessimism.

	Subj. Equity Premium	Subj. Term Premium	Size of Subj. Permanent Component
	$E[\tilde{E}_t(\log R^m_{t+1}/R^1_{t+1})]$	$E[\tilde{E}_t(\log R^{(k)}_{t+1}/R^1_{t+1})]$	$E[\tilde{L}_t(\Delta\tilde{H}_{t+1})/\tilde{L}_t(\Delta\tilde{S}_{t+1})]$
A. Nagel-Xu.			
$k = 10$ years	0.0476	-0.0175	0.8817
	(0.0044)	(0.0075)	(0.0547)
$k = 30$ years		-0.1454	0.9740
		(0.0137)	(0.0234)
$k = 10$ (backfill)		-0.0226	0.9098
		(0.0035)	(0.0294)
$k = 30$ (backfill)		-0.1499	0.9737
		(0.0161)	(0.0139)
B. Livingston.			
$k = 10$ years	0.0300	-0.0175	0.9080
	(0.0081)	(0.0075)	(0.0383)
$k = 30$ years		-0.1454	0.9719
		(0.0137)	(0.0271)
$k = 10$ (backfill)		-0.0226	0.9288
		(0.0035)	(0.0215)
$k = 30$ (backfill)		-0.1499	0.9478
		(0.0161)	(0.0260)
C. CFO.			
$k = 10$ years	0.0297	-0.0175	0.8686
	(0.0038)	(0.0075)	(0.0561)
$k = 30$ years		-0.1454	0.9848
		(0.0137)	(0.0149)
$k = 10$ (backfill)		-0.0226	0.8816
		(0.0035)	(0.0498)
$k = 30$ (backfill)		-0.1499	0.9627
		(0.0161)	(0.0254)

Table 2. Size of permanent risks in marginal utility. This table displays estimates of the subjective volatility bound from Corollary 2. Panels A, B, and C use the Nagel-Xu, Livingston, and CFO survey expectations for the market, respectively. The rows labeled "backfill" refer to BCFF forecasts fitted to the contemporaneous and lagged yield curve, and then projected to other periods in which the survey is not available. Newey-West asymptotic standard errors using 3 years of lags (i.e., 12 periods for Nagel-Xu, 6 periods for Livingston, 12 periods for CFO) are shown in parentheses.

the other extreme, it is possible that agents hold very distorted beliefs that fluctuate dramatically, in which case most of the permanent shocks are really shocks to beliefs rather than risks. This is mathematically coherent: belief distortions, being martingales, only feature permanent shocks. Under this belief-based explanation, the total amount of marginal utility variation is small, and so the fact that permanent risks drive this small variation is not economically important. Of course, the world probably lives somewhere in between these two polar opposite explanations. But where in between?

3.3 Main result: size of risks in the SDF

The remainder of this section attempts to answer the principal question of the paper: where does the world live on the spectrum between risk-based volatility and belief-distortion-based volatility? Table 3 presents the main empirical result that risk accounts for at least half the volatility of the SDF, with several estimates significantly higher. The table displays each unconditional mean in Proposition 1, followed by ratios of the subjective-to-objective bounds, as in Proposition 2. Standard errors assume the same Newey-West lag structure as before, with the standard errors for the volatility ratios computed using the Delta Method.

Strictly speaking, the subjective-to-objective ratio may not be meaningful if the stock market is far from the growth-optimal portfolio. But as Proposition 2 makes clear, if the stock market is closer to growth-optimal than it is to perceived growth-optimal, the ratios provide lower bounds on the contribution of risk to various components of the SDF (i.e., to the risk-neutral density $R^f \Delta S$ or the permanent component ΔH, respectively). A weaker interpretation is also permitted: if S represents an SDF that prices the stock market, riskless rate, and long bond (S may be taken to be unique by projecting into this space of returns), then all I require for the subjective-to-objective ratios to be meaningful is that the stock market is close to growth-optimal *among this subset of assets*.

The extreme pessimism in long bond forecasts, particularly the 30-year bond, results in enormous estimates of $L(\Delta\tilde{H})/L(\Delta H)$ in the final column.[9] Taking these estimates at face value, over 100% of the volatility of the permanent component of the SDF is risk-based. However, as mentioned earlier, such a large volatility ratio estimate may be due partly to the specificity of the data sample; the very large standard errors suggest that the large estimate may also be due to the tremendous volatility in long-term bond expectations. Still, suppose $L(\Delta\tilde{H})/L(\Delta H) > L(R^f\Delta\tilde{S})/L(R^f\Delta S)$ is true, even if the gap is smaller than Table 3 suggests, meaning H is driven by risks to a greater extent than is the overall SDF S. This happens if belief fluctuations partly manifest in transitory SDF variation, i.e., via G. The models in Sections 4-5 contain mechanisms of this type.

[9] While Proposition 2 shows that the lower bounds technically cannot be above 1, the final column of Table 3 is documenting estimates of the right-hand-side of (22) ignoring the minimum operator.

	Obj. Equity Premium $E[\log R^m / R^1]$	Subj. Equity Premium $E[\tilde{E}_t(\log R^m_{t+1} / R^1_{t+1})]$	Size of Risks in SDF $L(R^1 \Delta \tilde{S}) / L(R^1 \Delta S)$	Obj. Equity minus Term Premium $E[\log R^m / R^{(k)}]$	Subj. Equity minus Term Premium $E[\tilde{E}_t(\log R^m_{t+1} / R^{(k)}_{t+1})]$	Size of Risks in Permanent Component $L(\Delta \tilde{H}) / L(\Delta H)$
A. Nagel-Xu.						
$k = 10$ years	0.0555	0.0476	0.8567	0.0455	0.0651	1.4290
	(0.0161)	(0.0044)	(0.2449)	(0.0172)	(0.0081)	(0.4817)
$k = 30$ years				0.0464	0.1930	4.1568
				(0.0184)	(0.0138)	(1.6913)
B. Livingston.						
$k = 10$ years	0.0555	0.0300	0.5412	0.0455	0.0476	1.0443
	(0.0161)	(0.0081)	(0.1566)	(0.0172)	(0.0069)	(0.3165)
$k = 30$ years				0.0464	0.1755	3.7795
				(0.0184)	(0.0127)	(1.3275)
C. CFO.						
$k = 10$ years	0.0555	0.0297	0.5349	0.0455	0.0472	1.0367
	(0.0161)	(0.0038)	(0.1524)	(0.0172)	(0.0071)	(0.3365)
$k = 30$ years				0.0464	0.1751	3.7720
				(0.0184)	(0.0132)	(1.5321)

Table 3. Size of risks in the SDF. This table displays various estimates of the importance of risks versus belief volatility. Panels A, B, and C use the Nagel-Xu, Livingston, and CFO survey expectations for the market, respectively. Newey-West asymptotic standard errors using 3 years of lags (i.e., 12 periods for Nagel-Xu, 6 periods for Livingston, 12 periods for CFO) are shown in parentheses.

3.4 Robustness: optimizing the returns

As mentioned, my bounds are more likely to be valid if an asset is closer to objective growth-optimal, which may be a poor assumption for the aggregate stock market. I therefore undertake several robustness exercises below, including leverage, volatility-timing, and a kitchen-sink approach that accounts for a litany of factors.

The most obvious modification to the baseline trading strategy is to introduce a constant amount of leverage in order to optimize log returns. In particular, for any constant equity share θ, I construct the following return:

$$R^{\theta}_{t+1} := \theta R^{m}_{t+1} + (1-\theta) R^{f}_{t}, \quad \theta \text{ constant.} \tag{25}$$

Starting from (25), I estimate the objective expected excess log return $\mathbb{E}[\log(R^{\theta}_{t+1})] - \mathbb{E}[\log(R^{f}_{t})]$ for this strategy. Similarly, I can also use the surveys to compute proxies for the subjective version $\mathbb{E}[\tilde{\mathbb{E}}_t[\log(R^{\theta}_{t+1}) - \log(R^{f}_{t})]]$ using the same lognormal approximation as in the baseline results.[10]

Table 4 presents the results of these leveraged returns for various values of θ. Very similar to the main result, I continue to find that risk accounts for at least half the volatility of the SDF. Implementing leveraged versions of the market makes almost no difference to such results. The basic reason is that, while leverage does improve the objective return (the first column in Table 4), it also improves the subjective return series by a similar amount.

Next, I present two sets of very conservative bounds. These two exercises are similar in the sense that they "optimize" the portfolio behind the objective expected return series while making no changes to the portfolio for the subjective expected returns. This also makes them somewhat unfair comparisons that likely understate the importance of risk. The first exercise constructs the objective return series via a market-timing strategy in the spirit of "volatility-managed portfolios," a la Moreira and Muir (2017). The second exercise follows the methodology in Chen, Pelger and Zhu (2024b) to construct a proxy for the SDF, hence a proxy for the growth-optimal portfolio R^*.

The first exercise, following Moreira and Muir (2017), constructs a return series by increasing the portfolio's equity share when market variance is expected to be low. In

[10] Specifically, we have the approximation

$$\tilde{\mathbb{E}}_t[\log(R^{\theta}_{t+1})] \approx \log \tilde{\mathbb{E}}_t[R^{\theta}_{t+1}] - \frac{1}{2}\tilde{\mathrm{Var}}_t[R^{\theta}_{t+1}] = \log\left(\theta \tilde{\mathbb{E}}_t[R^{m}_{t+1}] + (1-\theta)R^{f}_{t}\right) - \frac{1}{2}\theta^2 \tilde{\mathrm{Var}}_t[R^{m}_{t+1}],$$

all of which are available in surveys. Then, to estimate $\mathbb{E}[\tilde{\mathbb{E}}_t[\log(R^{\theta}_{t+1}) - \log(R^{f}_{t})]]$, we take the time series average of the above quantity minus the time series average of $\log(R^{f}_{t})$.

equity share θ	Objective Premium $E[\log R^\theta / R^1]$	**A. Nagel-Xu.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$	**B. Livingston.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$	**C. CFO.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$
$\theta = 0.5$	0.0313	0.8408	0.5620	0.5501
	(0.0079)	(0.2114)	(0.1427)	(0.1384)
$\theta = 1$ (baseline)	0.0555	0.8567	0.5412	0.5349
	(0.0161)	(0.2449)	(0.1566)	(0.1524)
$\theta = 1.5$	0.0722	0.8821	0.5179	0.5178
	(0.0250)	(0.2995)	(0.1779)	(0.1743)
$\theta = 2$	0.0786	0.9513	0.5055	0.5131
	(0.0361)	(0.4268)	(0.2264)	(0.2267)

Table 4. Portfolios with constant leverage. This table displays estimates of the importance of risks versus belief volatility, using portfolios with constant leverage. The various rows indicate different equity shares θ in equation (25). Note that increasing θ above 2 is not possible in the time series, because a requirement to compute log returns is that $\theta R^m_{t+1} + (1-\theta)R^f_t > 0$, which is violated in some periods for $\theta > 2$. Panels A, B, and C use the Nagel-Xu, Livingston, and CFO survey expectations for the market, respectively. Newey-West asymptotic standard errors using 3 years of lags (i.e., 12 periods for Nagel-Xu, 6 periods for Livingston, 12 periods for CFO) are shown in parentheses.

particular, I construct the following return:

$$R^\theta_{t+1} := \theta_t R^m_{t+1} + (1-\theta_t)R^f_t, \quad \theta_t \propto \frac{1}{\text{variance measure at time } t}. \tag{26}$$

I take one of three measures for market variance. In the first approach, I use an AR(1) model to create $\hat{v}_{t+1}$, the month-t forecast of month-$t+1$ market variance—see Appendix E for more details. In the second approach, I simply use last month's realized variance v_t. The third approach uses VIX^2_t as the variance measure. To isolate the role of market timing and not introduce any average amount of leverage, I rescale portfolio weights so that the time-series average of θ_t is always one in each case. I also rescale θ_t so that it never exceeds 3 (i.e., this is the maximal allowable leverage).[11] One minor difference from the previous results is that these portfolios are rebalanced monthly, and so R^f_t is a one-month riskless rate here. I convert these monthly return series into annual returns by compounding.

The results, displayed in Table 5, are somewhat weaker than those so far. Rows 2-4 show that the the size of risks in the SDF is lower-bounded by between 28% and 60%, depending on the survey and market-timing strategy. This is because the volatility-managed portfolios perform very well, with average log excess returns between 7.9-

[11] This leverage constraint only affects the portfolio that is based on realized variance, which is highly volatile. The two portfolios based on the forecasted variance and the VIX-squared are unaffected.

equity share θ_t	Objective Premium $E[\log R^\theta / R^1]$	**A. Nagel-Xu.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$	**B. Livingston.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$	**C. CFO.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$
$\theta_t = 1$ (baseline)	0.0588	0.8093	0.5112	0.5053
	(0.0158)	(0.2127)	(0.1394)	(0.1332)
$\theta_t \propto 1/\hat{v}_{t+1}$	0.1087	0.4377	0.2765	0.2733
	(0.0166)	(0.0575)	(0.0475)	(0.0363)
$\theta_t \propto 1/v_t$	0.0794	0.5990	0.3784	0.3740
	(0.0163)	(0.1124)	(0.0822)	(0.0706)
$\theta_t \propto 1/\text{VIX}_t^2$	0.1084	0.4386	0.2771	0.2739
	(0.0209)	(0.0679)	(0.0511)	(0.0452)
$\theta_t \propto 1/\hat{v}_t$ (lag)	0.0593	0.8024	0.5068	0.5010
	(0.0153)	(0.2003)	(0.1317)	(0.1254)
$\theta_t \propto 1/v_{t-1}$ (lag)	0.0594	0.8002	0.5054	0.4996
	(0.0152)	(0.1982)	(0.1323)	(0.1241)
$\theta_t \propto 1/\text{VIX}_{t-1}^2$ (lag)	0.0929	0.5122	0.3235	0.3198
	(0.0219)	(0.0990)	(0.0692)	(0.0647)

Table 5. Portfolios with volatility market timing. This table displays estimates of the importance of risks versus belief volatility, using portfolios with time-varying equity shares. The various rows indicate different market-timing strategies θ_t in equation (26). In each strategy, I rescale portfolio weights so that the time-series average of θ_t is always one and so that $\max(\theta_t) = 3$. Panels A, B, and C use the Nagel-Xu, Livingston, and CFO survey expectations for the market, respectively. Newey-West asymptotic standard errors using 3 years of lags (i.e., 12 periods for Nagel-Xu, 6 periods for Livingston, 12 periods for CFO) are shown in parentheses.

10.9% per annum, depending on which strategy is adopted.[12] Of course, the reader should keep in mind that Table 5's comparison of objective versus subjective returns is a bit unfair: for each of the volatility-managed portfolios, the subjective counterpart is simply the baseline with $\theta_t = 1$ and so is not optimized in any way.

An additional concern with market-timing strategies is that they may be sensitive to the sample period, predictors used, and precise construction of the market-timing strategy. To illustrate this in a simple way here, the final three rows of Table 5 compute very similar market-timing strategies, except that the variance signal is lagged by one month. Although variances are generally quite persistent, this lagging procedure dramatically changes the results, with the estimates for the size of risk in the SDF now much closer to the baseline levels. (The portfolio using VIX as its variance signal is more resilient to the lagging of its signal than the others, but this further emphasizes that the choice of predictor variables matters delicately.)

The final exercise, following Chen et al. (2024b), uses a combination of neural networks and an adversarial estimation method to estimate jointly the SDF and relevant test assets—in particular, which weights to put on the cross-section of assets, and at

[12] The three volatility-managed portfolios have Sharpe ratios of 0.78, 0.58, and 0.69, respectively.

which times. The result is an SDF which takes the form $\frac{S_{t+1}}{S_t} = 1 - \sum_{n=1}^{N} \omega(I_t, I_{n,t}) R^e_{n,t+1}$, where $R^e_{n,t+1}$ is the excess return of stock n, and where ω are portfolio weights. These weights are estimated as a function of $(I_t, I_{n,t})$, where I_t are macroeconomic time-series factors and $I_{n,t}$ is a vector of characteristics for stock n. Market timing is captured by the extent to which ω depends on macro time-series factors in I_t, while stock-picking is captured by the way ω depends on characteristics in $I_{n,t}$. Given the assumed structure of the SDF, we construct the proxy for the growth-optimal portfolio by $R^*_{t+1} = \frac{S_t}{S_{t+1}} = (1 - \sum_{n=1}^{N} \omega(I_t, I_{n,t}) R^e_{n,t+1})^{-1}$.[13]

The first row of Table 6 shows that the bound does indeed become more conservative: in the baseline optimal SDF construction, the size of risks in the SDF is lower-bounded by between 31% and 50%. Again, this is because the baseline estimate for the growth-optimal portfolio performs extremely well, obtaining an average log excess return of 9.6% per annum.[14]

To place the subjective returns on more equal footing, I augment them with market timing as well. Given a survey-based estimate of one-year expected market returns $\tilde{\mathbb{E}}_t[R^m_{t+1}]$ and variance $\tilde{\mathrm{Var}}_t[R^m_{t+1}]$, I form the mean-variance optimal portfolio

$$\tilde{R}^*_{t+1} := \tilde{\theta}_t R^m_{t+1} + (1 - \tilde{\theta}_t) R^f_t, \quad \text{where} \quad \tilde{\theta}_t := \frac{\tilde{\mathbb{E}}_t[R^m_{t+1}] - R^f_t}{\tilde{\mathrm{Var}}_t[R^m_{t+1}]}. \tag{27}$$

This approximately delivers the subjective growth-optimal portfolio when restricting trade to the aggregate stock market and short-term bonds. The second row of Table 6 displays the results of the comparison between the baseline growth-optimal return R^* against this subjective version $\tilde{R}^*$. Under this fairer comparison, the results are much closer to our original results, with at least half of the SDF volatility being due to risk.

As is well known, such estimation methods using a rich cross-section of stocks and a rich set of time series predictors can potentially overstate the returns of the true growth-

[13]The results would be almost identical if we instead used $R^*_{t+1} = 1 + \sum_{n=1}^{N} \omega(I_t, I_{n,t}) R^e_{n,t+1}$, which more clearly lies in the return space if the riskless rate was equal to 1. While the actual riskless rate in the data differs from 1, note that the estimation procedure of Chen et al. (2024b) only uses excess returns and therefore does not correctly price the riskless rate. For that reason, it is consistent with the form of the SDF to assume a riskless rate of 1 in forming the growth-optimal portfolio. Nevertheless, we use the theoretically-correct specification of $R^*_{t+1} = S_t / S_{t+1}$. As a minor additional note, the estimation is done at the monthly level, so we annualize the returns of this growth-optimal portfolio for our tests.

[14]Or in arithmetic units, this portfolio has an average annual excess return of 10.8%, with a standard deviation of only 6.7%, implying an annualized Sharpe ratio of 1.61.

Notice that, despite the large cross-section of assets and large set of predictor variables, the optimized portfolio in Table 6 (with $\mathbb{E}[\log R^* / R^1] = 9.55\%$) performs slightly worse than the best volatility-managed portfolios in Table 5 (which had $\mathbb{E}[\log R^\theta / R^1] \approx 10.9\%$). Because the machine-learning approach estimates and tests its method based on out-of-sample performance, this suggests that the baseline volatility-managed portfolio features some amount of in-sample data mining.

	Objective Premium $E[\log R^*/R^1]$	**A. Nagel-Xu.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$	**B. Livingston.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$	**C. CFO.** Size of Risks in SDF $L(R^1\Delta\tilde{S})/L(R^1\Delta S)$
baseline SDF	0.0955	0.4981	0.3146	0.3110
	(0.0118)	(0.0456)	(0.0453)	(0.0244)
baseline SDF vs $\tilde{R}^*$	0.0955	0.8695	0.6911	0.4791
	(0.0118)	(0.0873)	(0.0664)	(0.0400)
top 30pct stocks	0.0508	0.9355	0.5909	0.5841
	(0.0100)	(0.1270)	(0.1050)	(0.0713)
top 20pct stocks	0.0450	1.0582	0.6684	0.6607
	(0.0099)	(0.1677)	(0.1278)	(0.0976)
top 10pct stocks	0.0370	1.2852	0.8118	0.8024
	(0.0104)	(0.2819)	(0.1932)	(0.1687)
no macro vars	0.0718	0.6622	0.4183	0.4134
	(0.0064)	(0.0539)	(0.0589)	(0.0248)

Table 6. Machine-learning-based growth-optimal portfolio. This table provides estimates of the importance of risks versus belief volatility, using proxies for the growth-optimal portfolio R^* from Chen et al. (2024b). The various rows indicate different specifications of the SDF-mimicking portfolio, or equivalently different growth-optimal portfolios R^*. The first row "baseline SDF" uses the baseline specification, data, and neural network settings in Chen et al. (2024b). The second row "baseline SDF vs $\tilde{R}^*$" computes the survey-based expected returns using the strategy $\tilde{R}^*$ defined in (27). The rows labelled "top XXpct stocks" uses only the top XX percent largest stocks to construct R^*. The row "no macro vars" does not use macroeconomic factors I_t to construct the portfolio weights for R^*. Panels A, B, and C use the Nagel-Xu, Livingston, and CFO survey expectations for the market, respectively. Newey-West asymptotic standard errors using 3 years of lags (i.e., 12 periods for Nagel-Xu, 6 periods for Livingston, 12 periods for CFO) are shown in parentheses.

optimal portfolio. Specifically, many apparently anomalous strategies rely heavily on small stocks, while market timing is very sensitive to the predictors used and the sample period. To explore the sensitivity of the method to these issues, I re-run the method by restricting the SDF in one of two ways: (i) disallowing the SDF to load on small stocks; and (ii) disallowing the SDF to depend on the macro factors.[15] As shown in the subsequent rows of Table 6, these SDFs imply a significantly-worse growth-optimal portfolio. This relates to the findings of Jensen et al. (2024), in which many types of optimal portfolios prescribed by machine-learning methods are shown to feature large amounts of turnover and transactions costs, hence poor net-of-cost performance. I don't directly take transactions costs into account, but one can think of these exercises—restricting the ability to trade small stocks and restricting extreme market timing—as ways of checking the robustness of the portfolio to such costs. As it relates to Table 6, the poor performance of the slightly perturbed growth-optimal portfolio means that the size of risks in

[15]For the small stock exclusion, the SDF is estimated including all stocks but then evaluated on only a subset of stocks, meaning that the ω weights are simply set to zero for those stocks. This is analogous to the robustness exercise Chen et al. (2024b) conduct in their Internet Appendix Table IA.VIII.

the SDF is once again very high. Given the difficulty in confidently pinning down the objective growth-optimal portfolio, I interpret these findings as showing that my main result—namely that risks constitute the majority of SDF volatility—is robust.

4 Example structural models with long-term risks

I provide three example models to help interpret the general theory above. All models are exchange economies but with differing endowment and belief structures. The first model features rational expectations about a time-varying long-run growth rate (Bansal and Yaron, 2004). The second model features distorted beliefs, formed using past growth extrapolation, about future growth (Collin-Dufresne et al., 2017; Nagel and Xu, 2022). In these two settings, there is a representative agent with recursive preferences a la Epstein and Zin (1989) and elasticity of intertemporal substitution (EIS) equal to one. These preferences are described by the utility recursion (Bellman equation)

$$U_t = (1-\beta)\log C_t + \beta \frac{\log \tilde{\mathbb{E}}_t[\exp((1-\gamma)U_{t+1})]}{1-\gamma}, \tag{28}$$

where $\tilde{\mathbb{E}}$ is the agent's subjective expectation, γ is her risk aversion, and β is her discount factor. The third model also features growth extrapolation but with CRRA preferences for comparison. After presenting the models, Section 4.4 compares them across various measures of SDF volatility. Appendix B contains some of the derivations.

4.1 Long-run risks

In this first example, the agent has rational expectations: $\tilde{\mathbb{E}} = \mathbb{E}$. There are two shocks which are jointly Normal, $W := (W^{(1)}, W^{(2)})' \sim N(0, I)$. Aggregate consumption follows

$$\log C_{t+1} = \log C_t + x_t + \sigma_c \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \cdot W_{t+1},$$

and its *expected growth rate* x_t has AR(1) dynamics

$$x_{t+1} = (1-\rho)\bar{x} + \rho x_t + \sigma_x \left(\begin{smallmatrix} (1-\kappa^2)^{1/2} \\ \kappa \end{smallmatrix}\right) \cdot W_{t+1} \tag{29}$$

The parameter ρ captures the persistence of expected consumption, and κ controls the correlation between consumption and growth shocks. A typical specification sets $\kappa = 1$, but this flexibility will be helpful going forward.

Solving the model, the value function is given by

$$U_t = \text{constant} + \log C_t + \frac{\beta}{1-\beta\rho} x_t$$

From this solution, we can write the SDF explicitly as

$$\frac{S_{t+1}}{S_t} = \beta \frac{C_t}{C_{t+1}} \frac{\exp[(1-\gamma)U_{t+1}]}{\mathbb{E}_t[\exp[(1-\gamma)U_{t+1}]]} = \exp\Big\{ -r_t - \pi \cdot W_{t+1} - \frac{1}{2}\|\pi\|^2 \Big\}, \tag{30}$$

where the one-period interest rate $r_t = \log R_t^f$ and market price of risk π are given by

$$r_t := -\log(\beta) + x_t + \frac{1}{2}\sigma_c^2 - \gamma\sigma_c^2 - \frac{(\gamma-1)\beta}{1-\beta\rho}(1-\kappa^2)^{1/2}\sigma_x\sigma_c \tag{31}$$

$$\pi := \gamma\sigma_c \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + \frac{(\gamma-1)\beta}{1-\beta\rho}\sigma_x \left(\begin{smallmatrix} (1-\kappa^2)^{1/2} \\ \kappa \end{smallmatrix}\right). \tag{32}$$

The permanent and transitory components of the SDF are given by

$$\frac{H_{t+1}}{H_t} = \exp\Big\{ -\frac{1}{2}\Big\|\pi + \frac{\sigma_x}{1-\rho}\left(\begin{smallmatrix} (1-\kappa^2)^{1/2} \\ \kappa \end{smallmatrix}\right)\Big\|^2 - \Big(\pi + \frac{\sigma_x}{1-\rho}\left(\begin{smallmatrix} (1-\kappa^2)^{1/2} \\ \kappa \end{smallmatrix}\right)\Big) \cdot W_{t+1} \Big\} \tag{33}$$

$$\frac{G_{t+1}}{G_t} = \exp\Big[\eta + \frac{1}{1-\rho}(x_{t+1} - x_t)\Big], \tag{34}$$

for some constant η. See Appendix B for these derivations. For the quantitative analysis below, I adopt a relatively typical calibration of this model, setting $\gamma = 10$, $\beta = 0.99$, $\rho = 0.95$, $\sigma_c = 0.02$, $\sigma_x = 0.002$, and $\kappa = 0.9$ (Bansal et al., 2012).

4.2 Growth extrapolation

Now, consider a model in which actual growth is IID but agents perceive a stochastic growth rate. It turns out, with the appropriate modeling of beliefs, we can just reinterpret the previous example as x representing the *subjective expected growth rate*, which remains constant at $\bar{x}$ under the rational expectation. This will be a model with long-run risks only in investors' heads.

Therefore, the solution will be identical to the actual long-run risks model but with a corresponding reinterpretation of the results. The expression in (30) now corresponds to investors' marginal utility $\tilde{S}_t$. Its permanent component $\tilde{H}_t$ is given by the expression in (33). In essence, the extrapolative growth model can be solved as a long-run risk model but then simulated by turning off long-run risks.

To justify these claims, consider the following enriched version of an extrapolative growth model. True consumption dynamics are

$$\log C_{t+1} = \log C_t + \bar{x} + \sigma_c \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \cdot W_{t+1}.$$

While conditional mean growth is constant at $\bar{x}$, agents are not convinced of this fact, and they do not observe the objective shocks W either. Building on many extrapolation models, agents will learn about the mean consumption growth via *constant-gain learning*. I will also allow the learning process to be subject to *sentiment shocks*.

Mathematically, let x_t denote the agent's perceived expected growth rate, i.e., $x_t := \tilde{\mathbb{E}}_t[\log C_{t+1} - \log C_t]$. Then, the agent's perception of consumption dynamics are

$$\log C_{t+1} = \log C_t + x_t + \sigma_c \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \cdot \tilde{W}_{t+1},$$

where $\tilde{W}$ represent perceived shocks, i.e., a normally-distributed variable which has zero mean under the agent's belief, $\tilde{\mathbb{E}}_t[\tilde{W}_{t+1}] = 0$. Learning is modeled as follows. The posterior expected growth is given by

$$x_t = (1-\phi)\bar{x} + \phi z_t, \tag{35}$$

where z_t is a "sentiment" variable that follows

$$z_{t+1} - z_t = \underbrace{\lambda[\log C_{t+1} - \log C_t - z_t]}_{\text{constant gain learning}} + \underbrace{\left(\begin{smallmatrix} \nu_1 - \lambda\sigma_c \\ \nu_2 \end{smallmatrix}\right) \cdot \tilde{W}_{t+1}}_{\text{sentiment shocks}} \tag{36}$$

In words, sentiment increases with excess growth realizations $\log C_{t+1} - \log C_t - z_t$, at a fixed learning rate λ, and is also subject to additional shocks. After sentiment is determined, posterior expected growth is determined by a type of "model averaging" between the correct time-invariant growth $\bar{x}$ and the growth sentiment z_t. The parameter ϕ governs the degree of rationality: $\phi = 0$ is fully rational, while $\phi = 1$ is fully extrapolative. Alternatively, as shown in Nagel and Xu (2022), $1-\phi$ captures the weight a constant-gain learner puts on his prior.

Combining equations (35)-(36) with the perceived consumption law of motion yields the belief evolution

$$x_{t+1} = \lambda(1-\phi)\bar{x} + (1 - \lambda(1-\phi))x_t + \phi\nu \cdot \tilde{W}_{t+1}, \tag{37}$$

where $\nu := (\nu_1, \nu_2)'$. From equation (37), we can see the analogy to the long-run risks model. By inspection, the dynamics in this equation are identical to those of equation (29), provided we set the prior dependence parameter to any interior value $\phi \in (0,1)$ and then put $\lambda = \frac{1-\rho}{1-\phi}$, $\nu_1 = \phi^{-1}(1-\kappa^2)^{1/2}\sigma_x$, and $\nu_2 = \phi^{-1}\kappa\sigma_x$.[16] In that case, the impact of λ is entirely through "perceived growth persistence" that I refer to going forward:

$$\tilde{\rho} := 1 - (1-\phi)\lambda.$$

Notice also that setting $\phi < 1$ serves a technical convenience by keeping $\tilde{\rho} < 1$ so that x_t is stationary under the subjective probability.

Given this isomorphism, the long-run risks solution above fully characterizes the investor marginal utility $\tilde{S}_t$ and its permanent component $\tilde{H}_t$. These are given by

$$\frac{\tilde{S}_{t+1}}{\tilde{S}_t} = \exp\Big\{ - r_t - \tilde{\pi} \cdot \tilde{W}_{t+1} - \frac{1}{2}\|\tilde{\pi}\|^2 \Big\} \tag{38}$$

$$\frac{\tilde{H}_{t+1}}{\tilde{H}_t} = \exp\Big\{ - \frac{1}{2}\Big\|\tilde{\pi} + \frac{\phi\nu}{1-\tilde{\rho}}\Big\|^2 - \Big(\tilde{\pi} + \frac{\phi\nu}{1-\tilde{\rho}}\Big) \cdot \tilde{W}_{t+1} \Big\} \tag{39}$$

where r_t and $\tilde{\pi}$ are the riskless rate and perceived risk prices, respectively,

$$r_t := -\log(\beta) + x_t + \frac{1}{2}\sigma_c^2 - \gamma\sigma_c^2 - \frac{(\gamma-1)\beta}{1-\beta\tilde{\rho}}\sigma_c\phi\nu_1$$

$$\tilde{\pi} := \gamma\sigma_c \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) + \frac{(\gamma-1)\beta}{1-\beta\tilde{\rho}}\phi\nu.$$

Similarly, the transitory component of the SDF, which recall is identical under both the subjective and objective probability, is given by

$$\frac{G_{t+1}}{G_t} = \exp\Big[\tilde{\eta} + \frac{1}{1-\tilde{\rho}}(x_{t+1} - x_t)\Big],$$

for some constant $\tilde{\eta}$. This fully solves the model from the perspective of agent's subjective probability.

It remains to characterize the belief distortion B_t, in order to map this solution back into the objective probability. In this conditionally lognormal environment, all belief distortions are characterized by a mean distortion to the shocks. In other words, while W are the true zero-mean shocks, the agent perceives them to have non-zero mean, $\tilde{\mathbb{E}}_t[W_{t+1}] = \mu_t$, and instead views $\tilde{W}_{t+1} = W_{t+1} - \mu_t$ as the relevant disturbance. This

[16] The special case without sentiment shocks ($\nu_1 = \lambda\sigma_c$ and $\nu_2 = 0$) would be like a long-run risk model with perfect correlation between level and growth shocks ($\kappa = 0$) and with a particular calibration of growth volatility ($\sigma_x = \phi\lambda\sigma_c$).

means that

$$\frac{B_{t+1}}{B_t} = \exp\left[\mu_t \cdot W_{t+1} - \frac{1}{2}\|\mu_t\|^2\right] \tag{40}$$

for some μ_t. To identify this mean, compare the true and perceived consumption dynamics to see that

$$\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \cdot \mu_t = \frac{x_t - \bar{x}}{\sigma_c}.$$

While this determines the first component of μ_t, the second component is not pinned down, because any value for it leads to the same perceived growth dynamics. Intuitively, pure sentiment shocks are not tied down by any objective dynamic. For theoretical clarity, let us set

$$\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) \cdot \mu_t = 0,$$

so that all results are generated via mechanisms that are tied down. (This can be easily generalized, and allowing a non-zero and possibly time-varying second component of μ_t would work like animal spirits, whereby agents can be optimistic/pessimistic about future sentiment shocks.)

Combining equations (38)-(39) with the belief distortion (40), we have the objective SDF and its permanent component

$$\frac{S_{t+1}}{S_t} = \exp\left\{ -r_t - (\tilde{\pi} - \mu_t) \cdot W_{t+1} - \frac{1}{2}\|\tilde{\pi} - \mu_t\|^2 \right\} \tag{41}$$

$$\frac{H_{t+1}}{H_t} = \exp\left\{ -\frac{1}{2}\left\|\tilde{\pi} - \mu_t + \frac{\phi\nu}{1-\tilde{\rho}}\right\|^2 - \left(\tilde{\pi} - \mu_t - \frac{\phi\nu}{1-\tilde{\rho}}\right) \cdot W_{t+1} \right\} \tag{42}$$

Thus, objective risk prices are lowered by μ_t relative to perceived risk prices. For example, if agents are pessimistic about growth, then the first component of μ_t is negative, which raises risk prices. From here, we may compute various volatility ratios. In the interest of space, these volatility ratios are derived and presented in Appendix B.

As a baseline, I adopt a standard calibration without sentiment shocks (Nagel and Xu, 2022). I set $\gamma = 4$, $\beta = 0.99$, $\sigma_c = 0.02$, $\phi = 0.9$ (almost fully extrapolative), $\lambda = 0.075$ (interpretable as roughly 13 years of effective information retention), and $\nu_1 - \lambda\sigma_c = \nu_2 = 0$ (no sentiment shocks).[17] This calibration generates perceived conditional growth

[17] Nagel and Xu (2022) set $\lambda = 0.018$ at a quarterly frequency, which is approximately 4×0.018 annually. Next, in their model $1 - \phi$ can be interpreted as the weight agents place on their prior. Their baseline sets

rate volatility $\phi\|\nu\| = 0.0013$ and a perceived persistence $\tilde{\rho} = 0.993$. Relative to the long-run risk model calibration, the extrapolative calibration features a lower short-run volatility of perceived growth but a higher persistence hence higher long-run volatility.

4.3 Growth extrapolation (CRRA)

For comparison, I also consider a version of the previous extrapolative growth model under CRRA utility with risk aversion γ. The sentiment and belief processes are identical to the previous setting. But now, marginal utility growth is simply

$$\frac{\tilde{S}_{t+1}}{\tilde{S}_t} = \beta\Big(\frac{C_{t+1}}{C_t}\Big)^{-\gamma} = \exp\Big\{ - r_t - \frac{1}{2}\|\tilde{\pi}\|^2 - \tilde{\pi}\cdot\tilde{W}_{t+1}\Big\},$$

where the riskless rate and perceived risk prices are the familiar expressions

$$\begin{aligned} r_t &:= -\log(\beta) + \gamma x_t - \frac{1}{2}(\gamma\sigma_c)^2 \\ \tilde{\pi} &:= \gamma\sigma_c\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \end{aligned}$$

Notice that the structure of the solution is very similar, with two key differences. First, perceived short-run risk prices $\tilde{\pi}$ do not inherit any exposure to perceived growth. Second, the perceived growth rate volatility enters instead into magnified interest rate volatility. The result of these two differences will be (i) much less marginal utility volatility, hence much more of SDF volatility attributable to beliefs; and (ii) much more of SDF volatility coming from transitory fluctuations. The expressions for the various components of the SDF are contained in Appendix B. We calibrate this model exactly as the Epstein-Zin version in the previous subsection, with the obvious exception that the EIS with CRRA utility is γ^{-1}, the inverse of the risk aversion parameter.

4.4 Comparing the models

Table 7 compares the three models in terms of various volatilities and volatility ratios.

Start with the long-run risk model (column 1). This is a model that can generate a large amount of SDF volatility via a very prominent permanent component. Indeed, amount of SDF volatility is $\sqrt{2L(\Delta S)} = \|\pi\| = 0.43$. As argued by Hansen and Jagannathan (1991), this large SDF volatility (i.e., large short-term risk prices or Sharpe ratios)

$\phi = 1$, but this would induce a belief that $\tilde{x}_t$ follows a random walk, which causes some technical issues for unconditional subjective volatilities. Thus, I pick a value ϕ close to 1 but slightly below.

	Long-Run Risk	Extrapolation (EZ)	Extrapolation (CRRA)
Annual Avg SDF Volatility			
$E[2L_t(\Delta S_{t+1})]^{1/2}$	0.4265	0.3574	0.1948
Size of Permanent Component			
$L(\Delta H)/L(\Delta S)$	1.1786	2.1275	17.5983
Size of Transitory Component			
$L(\Delta G)/L(\Delta S)$	0.0090	0.2635	14.1140
Risk Component			
$L(\Delta \tilde{S})/L(\Delta S)$	1.0000	0.7736	0.1677
Permanent Risk Component			
$L(\Delta \tilde{H})/L(\Delta H)$	1.0000	0.9118	0.9831
Belief Component			
$L(\Delta B)/L(\Delta S)$	0.0000	0.2471	0.8270

Table 7. Volatilities and volatility ratios. This table displays entropies and entropy ratios across the three example models, in their benchmark calibrations. The "Long-Run Risk" model has $\gamma = 10$, $\beta = 0.99$, $\rho = 0.95$, $\sigma_c = 0.02$, $\sigma_x = 0.002$, $\kappa = 0.9$. The "Extrapolation (EZ)" model has $\gamma = 4$, $\beta = 0.99$, $\phi = 0.9$, $\lambda = 0.075$, $\sigma_c = 0.02$, and $\nu = (\lambda\sigma_c, 0)'$. "Extrapolation (CRRA)" uses the same parameters as Extrapolation (EZ), except with CRRA preferences.

is an important target for asset-pricing models. At the same time, the size of the permanent and transitory components are $L(\Delta H)/L(\Delta S) = 1.18$ and $L(\Delta G)/L(\Delta S) = 0.01$, respectively. As argued by Alvarez and Jermann (2005), the permanent component must be large and the transitory component small. This class of models obtains all these features via a highly persistent growth rate (high ρ), combined with an investor who is both patient (high β) and risk averse (high γ).

Recall that the growth extrapolation model with EZ preferences (column 2) is very much like a long-run risks model but where the growth rate fluctuates merely in investors' heads. Therefore, it should not be surprising that this model succeeds in generating high SDF volatility, a large permanent component, and a modest transitory component. Because of the relatively high calibrated persistence of perceived growth ($\tilde{\rho} = 0.993$), the permanent component is even larger here than in the long-run risk calibration, $L(\Delta H)/L(\Delta S) = 2.13$. It turns out the transitory component is also larger, i.e., $L(\Delta G)/L(\Delta S) = 0.26$, because of negative correlation between G and H. In particular, optimism about perceived growth (high x_t) raises the transitory component of marginal utility, while simultaneously reducing the permanent component.

Perhaps more surprisingly, this EZ growth extrapolation model generates sizeable risk-based volatility. In fact, $L(\Delta \tilde{S})/L(\Delta S) = 0.7736$ and $L(\Delta \tilde{H})/L(\Delta H) = 0.9118$ are close to the empirical estimates in Table 3. Interestingly, only 24.71% of total SDF volatility comes directly from belief distortions in this model. This is because the action in this model, leading to the large value of $L(\Delta S)$, arises via an *interaction* between preferences and perceived growth dynamics. While shutting down the belief dynamics mitigates SDF volatility, so does eliminating the preference for early resolution of uncertainty. In general, when SDF volatility is created by interaction effects, it is not a priori clear how to attribute it; my framework provides a clear-cut answer to that question. Overall, this model thus generates significant SDF volatility, and other desirable properties, because marginal utility and its permanent component are sensitive to perceived growth shocks I refer to this mechanism as *subjective risk*, because beliefs indirectly create large risk.

By comparing the growth extrapolation models with EZ versus CRRA preferences, we see the criticality of subjective risk. In contrast to the EZ model, the CRRA model does not have any mechanism whereby beliefs fluctuations can impact marginal utility, which is purely a function of consumption. That is why most of its SDF volatility stems from beliefs rather than risks: the CRRA model features $L(\Delta\tilde{S})/L(\Delta S) = 0.1677$ and $L(\Delta B)/L(\Delta S) = 0.8270$, in stark contrast to my empirical estimates. It turns out that this counterfactually large direct role for beliefs is tightly related to another failure of this CRRA model—it features far too much transitory SDF volatility, with $L(\Delta G)/L(\Delta S) = 14.1$. This connection between direct belief-based volatility and transitory SDF volatility is further highlighted by the next exercise.

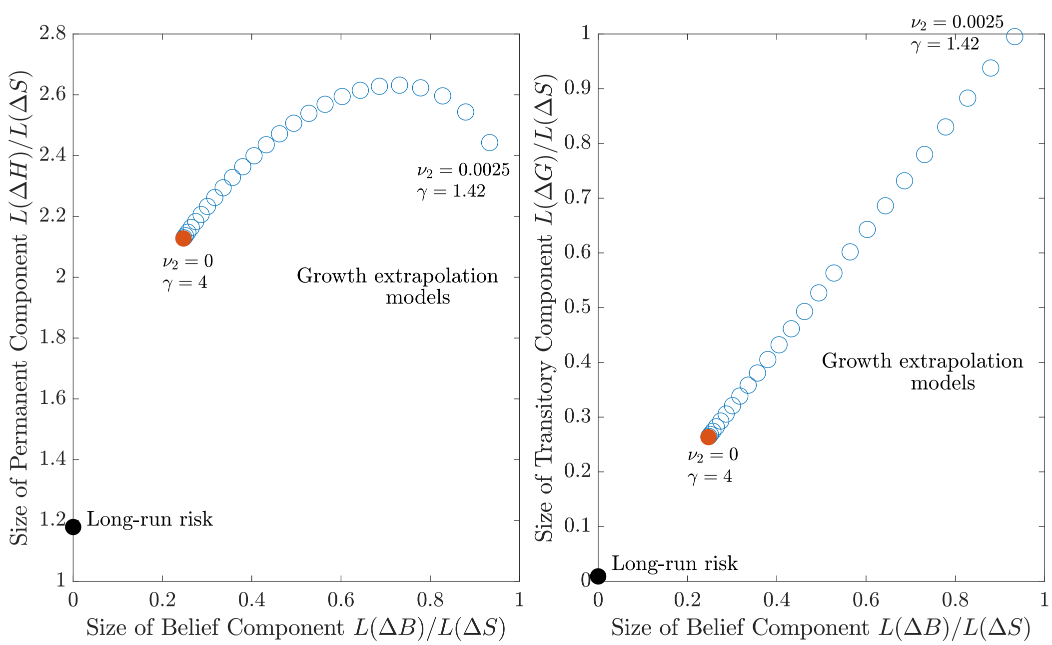

Figure 5. Importance of beliefs in various models. This table shows a comparison between the baseline long-run risk model (solid black dot), the baseline growth extrapolation model (solid red dot), and various growth extrapolation models with sentiment shocks (hollow blue circles), i.e., models with $\nu_2 > 0$. As ν_2 is increased, γ is also decreased to keep the measure of one-year average SDF volatility $\mathbb{E}[2L_t(\Delta S_{t+1})]^{1/2}$ unchanged at 0.3574 (the baseline value for the extrapolation model). The calibrations of the baseline models are the same as in Table 7.

In a final exercise, I add sentiment shocks to the extrapolation model, as a way to increase the volatility of belief distortions. The question is how this extra belief volatility affects the contributions to the SDF of risk, beliefs, and permanent/transitory shocks. Specifically, the extrapolation model is solved for different levels of ν_2 (the sensitivity of perceived growth to a shock orthogonal to consumption). To keep these various settings on a level playing field, I recalibrate risk aversion γ to hold fixed the model-implied average SDF volatility $\mathbb{E}[2L_t(\Delta S_{t+1})]^{1/2}$. That is, since sentiment shocks increase SDF volatility, I reduce γ as I increase ν_2. Figure 5 displays results for the size of the permanent component $L(\Delta H)$, transitory component $L(\Delta G)$, and belief component $L(\Delta B)$, all scaled by the total volatility of the SDF $L(\Delta S)$. The figure shows that sentiment shocks raise the contribution of beliefs to SDF volatility, as expected, but they do so by increasing the size of both the permanent and transitory components of the SDF. In this class of models, there is a tension: attributing large amounts of SDF volatility to beliefs necessitates a counterfactually transitory SDF. Thus, if distorted beliefs are to be taken seriously, their *direct impact* on SDF volatility must not be too large, and their impact must instead come *indirectly via subjective risk*.

To demonstrate that this conclusion is not an artifact of the particular calibration, I provide a theoretical result formalizing the tension between adding belief-based volatility while keeping transitory volatility low. Thus, the conclusion that SDF volatility must come via the subjective risk channel is somewhat general to this class of models.

Proposition 3. *Consider the growth extrapolation model with Epstein-Zin utility. For each parameter $\theta \in \{\gamma, \beta, \nu_2\}$, we have that $L(\Delta G)/L(\Delta S)$ and $L(\Delta B)/L(\Delta S)$ respond in the same direction to θ, i.e.,*

$$sign\Big(\frac{\partial}{\partial\theta}\frac{L(\Delta G)}{L(\Delta S)}\Big) = sign\Big(\frac{\partial}{\partial\theta}\frac{L(\Delta B)}{L(\Delta S)}\Big).$$

Furthermore, assuming either ϕ or $\tilde{\rho}$ are small enough, then the same conclusion holds for $\theta = \phi$.

Proposition 3 extends beyond comparative statics to γ and ν_2, as in the numerical example above, to β and ϕ as well. These preference and sentiment parameters are somewhat difficult to pin down directly without asset prices, and so Proposition 3 can also be thought of as a *partial identification* result: certain parts of the parameter space—such as low γ, high ν_2, low β, low ϕ, or low $\tilde{\rho}$—are ruled out by the requirement that we simultaneously match modest values of $L(\Delta B)/L(\Delta S)$ and $L(\Delta G)/L(\Delta S)$.

5 Example structural models with belief heterogeneity

In this section, I provide an alternative class of example models with heterogeneity in beliefs. I have two objectives. First, I want to illustrate how to aggregate beliefs and marginal utilities in this world to correspond to what is measured in the survey. This also maps nicely into surveys that contain relatively modest biases in the "consensus belief" but lots of disagreement. Second, the model helps reinforce the key finding that subjective risks are critically important drivers of SDF volatility. There is a large literature on belief disagreement and asset prices, which I do not summarize here. For now, just note that my model has two agent types who disagree about growth rates, experience sentiment shocks, have identical CRRA utility, trade in dynamically complete markets, and live on an infinite horizon. The closest paper is Dumas et al. (2009).[18]

Now to the details. I present this model in continuous time, because it adds significant tractability. The aggregate endowment follows a geometric Brownian motion,

$$\frac{dC_t}{C_t} = \mu_c dt + \sigma_c \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \cdot dW_t,$$

where W is a two-dimensional Brownian motion. The first shock drives aggregate consumption, while the second shock will drive beliefs directly and is independent of consumption growth. By no-arbitrage, the objective SDF follows

$$\frac{dS_t}{S_t} = -r_t dt - \pi_t \cdot dW_t,$$

for an endogenous interest rate r_t and risk price vector π_t.

For simplicity, I assume only two types of agents, indexed by $\theta \in \{0, 1\}$. Agents of type $\theta = 1$ hold pessimistic views about expected growth. By contrast, the beliefs of agents of type $\theta = 0$ are fully rational: she knows that the expected growth rate is the constant μ_c and observes all shocks correctly.

We nest both agent types and their beliefs as follows. Define agent θ's conditional expected growth rate $\mu_c^\theta := \frac{1}{dt}\mathbb{E}_t^\theta[dC_t/C_t] = \mu_c - \theta\sigma_c\beta$, with $\beta > 0$ capturing the degree

[18] A partial list of papers studying heterogeneous beliefs and asset prices are Detemple and Murthy (1994), Zapatero (1998), Basak (2000), Scheinkman and Xiong (2003), Buraschi and Jiltsov (2006), Kogan et al. (2006), Yan (2008), David (2008), Dumas et al. (2009), Bhamra and Uppal (2014), Kasa et al. (2014), Barberis et al. (2015), Chabakauri (2015), Atmaz and Basak (2018), Ehling et al. (2018), Kozak et al. (2018), Borovička (2020), Martin and Papadimitriou (2022), and Khorrami (2023).

of pessimism bias. That is, agent θ believes that

$$\frac{dC_t}{C_t} = \mu_c^\theta dt + \sigma_c \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \cdot d\tilde{W}_t^\theta,$$

where $\tilde{W}_t^\theta$ is a Brownian motion under agent θ's belief. To introduce sentiment risk in a way that economizes on state variables, I assume furthermore that the biases with regard to the second shock (sentiment) are proportional to that of the first shock (consumption). Given this setting, Girsanov's theorem then implies that the perceived shocks are characterized by

$$d\tilde{W}_t^\theta := dW_t + \theta\beta \left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right) dt \tag{43}$$

for some parameter χ that captures the volatility of sentiment shocks. Since sentiment shocks are orthogonal to all fundamentals, the parameter χ only serves to create trade but can influence endogenous prices through this mechanism. Summarizing this information, the martingale belief distortion for agent θ is given by

$$B_t^\theta = \exp\left[-\theta\beta \left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right) \cdot W_t - \frac{1}{2}(\theta\beta)^2(1+\chi^2)t \right]. \tag{44}$$

The assumption that agents' beliefs about growth are constant over time is a simplification. Given we have already studied some consequences of time-varying perceived growth in Section 4, the present assumption of constant perceived growth focuses the discussion here on novel issues.

As is well known, in most but not all models of belief heterogeneity, the "market selection hypothesis" holds, namely the long-run equilibrium is dominated by agents with the most accurate beliefs (Kogan et al., 2006; Yan, 2008; Dumas et al., 2009; Borovička, 2020). To avoid this issue and generically ensure survival of all agents, I also embed an overlapping generations structure as in Gârleanu and Panageas (2015) and Panageas (2020). In particular, I assume that agents die idiosyncratically at rate δ and are replaced by a mass δ of newborns arriving to the economy. The time-t population of surviving agents that were born at time $u \leq t$ is thus $\delta e^{-\delta(t-u)}$, while the total population is constant at $\int_{-\infty}^t \delta e^{-\delta(t-u)} du = 1$. I assume that η fraction of newborns pessimists; sometimes, I will write $\eta^1 := \eta$ and $\eta^0 := 1-\eta$ for the population shares of type-θ agents. To keep financial markets dynamically complete, I follow the standard practice and allow all agents to trade fairly priced annuity contracts to insure their death shocks—agents will optimally fully insure in these markets. Finally, to ensure that all agents start with some wealth, I assume an agent born at time u inherits a tree whose time $t \geq u$ dividends are given by $C_{t,u} := \kappa e^{-\kappa(t-u)} C_t$. The time-$t$ dividends of all trees ever born add up to the aggregate endowment, since $\int_{-\infty}^{t} C_{t,u} du = \int_{-\infty}^{t} \kappa e^{-\kappa(t-u)} C_t du = C_t$.

All agents have CRRA utility with common risk aversion γ and common discount rate ρ. Let $c_{u,t}^{\theta}$ be the time-t consumption flow for a type-θ agent born at time $u \leq t$. Given the details above, agents solve

$$\max_{(c_{t,u}^{\theta})_{u \geq t}} \mathbb{E}\Big[\int_t^{\infty} e^{-(\rho+\delta)(u-t)} \frac{B_u^{\theta}}{B_t^{\theta}} \frac{(c_{u,t}^{\theta})^{1-\gamma}}{1-\gamma} du \Big] \tag{45}$$

subject to the lifetime budget constraint

$$\mathbb{E}\Big[\int_t^{\infty} e^{-\delta(u-t)} \frac{S_u}{S_t} c_{u,t}^{\theta} du \Big] = \frac{P_{t,t}}{\delta}, \tag{46}$$

where $P_{t,t} = \mathbb{E}_t\big[\int_t^{\infty} \frac{S_u}{S_t} C_{u,t} du \big] = \mathbb{E}_t\big[\int_t^{\infty} \kappa e^{-\kappa(u-t)} \frac{S_u}{S_t} C_u du \big]$ denotes the (endogenous) time-t value of the trees born at time t, and the division by δ converts this into per-capita units. Notice that the death rate δ only enters the optimization problem as an augmented discount rate. Also notice that belief distortions are simply captured by B_t^{θ} inside the expectation of the utility function—this converts the agent's subjective expectation $\mathbb{E}^{\theta}$ to the objective expectation $\mathbb{E}$.

The equilibrium market clearing conditions are as follows. Define the aggregate consumption of type-θ agents, $c_t^{\theta} := \eta^{\theta} \int_{-\infty}^{t} \delta e^{-\delta(u-t)} c_{u,t}^{\theta} du$. The goods market clearing condition says $\sum_{\theta} c_t^{\theta} = C_t$. Rewrite this in terms of consumption shares by defining $x_t^{\theta} := c_t^{\theta}/C_t$ to be the consumption share of the type-θ agents and then noting that $\sum_{\theta} x_t^{\theta} = 1$. Additionally, the asset market clearing condition says that $\sum_{\theta} n_t^{\theta} = P_t$, where n_t^{θ} is the total net worth of type-θ agents and P_t is the total wealth in the economy, namely the total present value of all trees alive at time t. Because of the exponential decay of dividends, total wealth is simply equal to $P_t = \mathbb{E}_t\big[\int_t^{\infty} e^{-\kappa u} \frac{S_u}{S_t} C_u du \big] = P_{t,t}/\kappa$.[19]

I characterize the entire equilibrium in terms of a single state variable $x_t := x_t^1$, the consumption share of the pessimists. The dynamics of x_t are endogenous. Write

$$dx_t = \mu_{x,t} dt + \sigma_{x,t} \cdot dW_t$$

[19]Recall that the time-u dividends for trees born at time t are given by $\kappa e^{-\kappa(u-t)} C_u$. Hence, a time-$t$ claim to all the presently alive trees delivers total time-u dividends equal to $\int_{-\infty}^{t} \kappa e^{-\kappa(u-t)} C_u dt = e^{-\kappa(u-t)} C_u$, i.e., we sum over all trees born before t. The present-value of these dividends is the aggregate wealth, so $P_t := \mathbb{E}_t\big[\int_t^{\infty} e^{-\kappa u} \frac{S_u}{S_t} C_u du \big]$.

for some drift $\mu_{x,t}$ and diffusion $\sigma_{x,t}$ to be determined. As shown in Proposition C.1, this model can be solved up to a set of two ODEs, that are given in Appendix C. Proposition C.1 shows that equilibrium risk prices are given by

$$\pi_t = \gamma\sigma_c\binom{1}{0} + x_t\beta\binom{1}{\chi}. \tag{47}$$

While the pessimists' consumption share x_t is endogenous, we see that both larger pessimism (higher β) and more sentiment shocks (higher χ) both work to raise risk prices.

Given the SDF process, I obtain its permanent-transitory factorization. In this model, since the $\theta = 0$ agents are rational, their marginal utility is equated to the SDF, and so

$$\frac{S_t}{S_0} = e^{-\rho t}\Big(\frac{1-x_t}{1-x_0}\Big)^{-\gamma}\Big(\frac{C_t}{C_0}\Big)^{-\gamma}. \tag{48}$$

Since the state variable x_t is stationary, due to the OLG structure, the permanent component of S_t is precisely the permanent component of $C_t^{-\gamma}$.

The next question is how one should average across the heterogeneous agents in this model. To maintain consistency with the surveys, which obtain a "consensus expectation" by averaging across respondents, I define the consensus belief by the population-weighted average belief. As shown in Appendix D, this type of consensus belief preserves all the volatility bounds in this paper, so long as we interpret B as the cross-sectional average belief distortion, $\tilde{S}$ as the belief-weighted-average marginal utility, and $\tilde{H}$ as its permanent component. In other words, we first define B_t by

$$\frac{dB_t}{B_t} = \sum_\theta \eta^\theta \frac{dB_t^\theta}{B_t^\theta}. \tag{49}$$

Given our SDF S and its permanent component H, one can then define $\tilde{S}_t := S_t/B_t$ and $\tilde{H}_t := H_t/B_t$. At this point, we have all the various components of the SDF: S, H, G, B, $\tilde{S}$, and $\tilde{H}$. Appendix C presents the calculations of these components.

To compute various volatilities in this continuous-time model, I use "localized" versions of the entropy statistics. Localized entropies deliver clean expressions that permit analytical results as well. For a positive process X, define the local volatility operator

$$\mathcal{L}_t(X) := \lim_{\Delta\to 0} \Delta^{-1} L_t(X_{t+\Delta}). \tag{50}$$

Given the environment above has Brownian shocks, this localized volatility simply gives the quadratic variation of $\log X$, i.e., $\mathcal{L}_t(X) = \frac{1}{dt} d[\log X, W]_t$ (see Appendix A.6). To average over all states, I then compute unconditional versions $\mathbb{E}[\mathcal{L}_t(X)]$. I will compute these average local volatilities for the SDF, its permanent and transitory components, the belief distortion, and marginal utilities.

Table 8 presents these volatilities and their ratios in the model. The first column adopts a baseline calibration where pessimists have 3% lower expectations about output growth ($\beta = 0.03$). This calibration features equal population shares of rational agents and pessimists ($\eta = 0.5$), a moderate amount of sentiment shocks ($\chi = 5$), and typical level of risk aversion ($\gamma = 10$). The calibration produces volatilities that are roughly in line with the data: an annualized Sharpe ratio of 0.32, a permanent component that explains most of SDF volatility, and a dominant role for risk as opposed to beliefs.

	Baseline	More sentiment shocks	More pessimism
Annual Avg SDF Volatility			
$E[2\mathcal{L}_t(S)]^{1/2}$	0.3229	0.3305	0.3228
Size of Permanent Component			
$E[\mathcal{L}_t(H)]/E[\mathcal{L}_t(S)]$	0.8633	0.5272	0.5528
Size of Transitory Component			
$E[\mathcal{L}_t(G)]/E[\mathcal{L}_t(S)]$	0.0529	0.4260	0.2907
Risk Component			
$E[\mathcal{L}_t(\tilde{S})]/E[\mathcal{L}_t(S)]$	0.8609	0.5860	0.5370
Permanent Risk Component			
$E[\mathcal{L}_t(\tilde{H})]/E[\mathcal{L}_t(H)]$	0.9650	2.4414	1.3889
Belief Component			
$E[\mathcal{L}_t(B)]/E[\mathcal{L}_t(S)]$	0.0561	0.8258	0.3993

Table 8. Volatilities and volatility ratios (localized). This table displays entropies and entropy ratios in the disagreement model. The "baseline" calibration is $\gamma = 10$, $\rho = 0.01$, $\mu_c = 0.02$, $\sigma_c = 0.03$, $\delta = 0.02$, $\kappa = 0.033$, $\beta = 0.03$, $\chi = 5$, $\eta = 0.5$. The column "more sentiment shocks" uses the same parameters except for $\gamma = 8$ and $\chi = 20$. The column "more pessimism" uses the same parameters except for $\gamma = 8$ and $\beta = 0.08$. The local entropies are calculated from a 5,000 year simulation of the equilibrium.

Interestingly, despite the fact that a non-trivial share of agents are quite pessimistic, the direct role of belief distortions in SDF volatility is quite small at only 5.6%. This occurs for two reasons. First, remember that the "consensus belief" is defined as the cross-sectional average, which is in between the pessimistic and rational agents. Second, in contrast to a direct role, belief distortions also play an indirect role here: *disagreement creates risk*. Disagreement risk arises because rational agents must hold whatever assets the pessimistic agents sell them, and this requires equilibrium risk compensation. While the precise mechanism is different, disagreement risk here thus parallels the subjective risk channel from Section 4.

In this class of models, disagreement risk must be the key role played by belief heterogeneity. To demonstrate this, the second and third columns of Table 8 consider alternative calibrations with larger sentiment shocks (raise χ from 5 to 20) or more pessimism (raise β from 0.03 to 0.08). With larger sentiment shocks or more pessimism, agents trade more in a market that is orthogonal to aggregate consumption. While this trade does increase SDF volatility, it does so almost purely via belief volatility itself. To make a fair comparison across calibrations, I also recalibrate risk aversion to keep the overall level of Sharpe ratios roughly constant as I increase χ or increase β (I reduce γ from 10 to 8). The results show that large sentiment shocks or extreme disagreement raise the direct role of beliefs, reduce the risk-based component of the SDF, and reduce the persistence of the SDF. These results—a large direct role for beliefs and a large role for transitory shocks—are at odds with the data.

I complement these results with an analytical calculation. This calculation examines the conditional local entropies $\mathcal{L}_t$ and therefore ignores the (quantitatively minor) effects of a parameter on the stationary consumption distribution. As suggested by the calibration results, one can show that sentiment shocks always amplify the belief component and the transitory component of the SDF, whereas risk aversion always mitigates both components. Surprisingly, a similar story holds for every parameter: anything that increases the belief component $\mathcal{L}_t(B)$ also increases the transitory component $\mathcal{L}_t(G)$, and vice versa. This reinforces my claim that beliefs must impact equilibrium largely through risk, because otherwise they create too much transitory volatility in the SDF.

Proposition 4. *In the heterogeneous belief model, we have*

$$\mathcal{L}_t(G) = \left(\frac{x_t}{\eta}\right)^2 \mathcal{L}_t(B), \tag{51}$$

so $\mathcal{L}_t(G)$ and $\mathcal{L}_t(B)$ respond in the same direction to any parameter besides η.

6 Conclusion

Undoubtedly, asset markets are influenced jointly by risk, risk attitudes, and investors' subjective beliefs. For example, risk-based stories emphasize long-term risks, cyclical risk tolerance, balance-sheet constraints, and so on. On the other hand, pervasive violations of full-information rational expectations in macro-finance surveys suggest that belief dynamics must also matter. In the theoretical literature, rigorous frameworks alternatively emphasizing risks or beliefs can account for a variety of facts in asset markets.

Against this backdrop, this paper seeks to quantify *how much* of asset market dynamics should be attributable to risk versus beliefs. I find that risk is likely to be a more important factor. In particular, combining a volatility bounds approach with financial survey data in stocks and bonds, I estimate that at least 50% of SDF volatility must stem from risk. Some estimates for this risk-based fraction exceed 100%.

How should researchers proceed knowing that SDF volatility is primarily driven by marginal utility? One option is to abandon non-rational beliefs for the purpose of explaining aggregate asset-pricing puzzles. This approach would give up on matching the minutiae in financial surveys, taking the stance that belief distortions are more of a cumbersome distraction than a productive solution. Perfectly consistent with this perspective on aggregate asset pricing is a popular view that subjective beliefs about idiosyncratic components of cash flows create "excess volatility." This paper is silent on the importance of belief distortions about asset-specific cash flows.

A second option, that I elaborate on in this paper, is to refine the way in which beliefs can matter. Using some example structural models, I show how belief distortions must matter primarily by creating marginal utility volatility. What does this mean? In the growth extrapolation framework, marginal utility inherits important volatility from *subjective risk*. In the heterogeneous belief framework, marginal utility volatility is amplified by *disagreement-induced risk*. Having isolated these channels, further investigation into belief-induced risks is needed to bring additional nuance and clarity to a behavioral finance view of asset markets.

References

Adam, Klaus, Dmitry Matveev, and Stefan Nagel, "Do survey expectations of stock returns reflect risk adjustments?," *Journal of Monetary Economics*, 2021, *117*, 723–740.

Alvarez, Fernando and Urban J Jermann, "Using asset prices to measure the persistence of the marginal utility of wealth," *Econometrica*, 2005, *73* (6), 1977–2016.

Amromin, Gene and Steven A Sharpe, "From the horse's mouth: Economic conditions and investor expectations of risk and return," *Management Science*, 2014, *60* (4), 845–866.

Atmaz, Adem and Suleyman Basak, "Belief dispersion in the stock market," *The Journal of Finance*, 2018, *73* (3), 1225–1279.

Backus, David, Mikhail Chernov, and Ian Martin, "Disasters implied by equity index options," *The Journal of Finance*, 2011, *66* (6), 1969–2012.

_ , _ , and Stanley Zin, "Sources of entropy in representative agent models," *The Journal of Finance*, 2014, *69* (1), 51–99.

Bakshi, Gurdip and Fousseni Chabi-Yo, "Variance bounds on the permanent and transitory components of stochastic discount factors," *Journal of Financial Economics*, 2012, *105* (1), 191–208.

Bansal, Ravi and Amir Yaron, "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 2004, *59* (4), 1481–1509.

— **and Bruce N Lehmann**, "Growth-optimal portfolio restrictions on asset pricing models," *Macroeconomic dynamics*, 1997, *1* (2), 333–354.

— **, Dana Kiku, and Amir Yaron**, "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 2012, *1* (1), 183–221.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, "X-CAPM: An extrapolative capital asset pricing model," *Journal of Financial Economics*, 2015, *115* (1), 1–24.

Basak, Suleyman, "A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk," *Journal of Economic Dynamics and Control*, 2000, *24* (1), 63–95.

Ben-David, Itzhak, John R Graham, and Campbell R Harvey, "Managerial miscalibration," *The Quarterly journal of economics*, 2013, *128* (4), 1547–1584.

Bhamra, Harjoat S and Raman Uppal, "Asset prices with heterogeneity in preferences and beliefs," *The Review of Financial Studies*, 2014, *27* (2), 519–580.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, "Survey data and subjective beliefs in business cycle models," *Review of Economic Studies*, 2025, *92* (3), 1375–1437.

Bidder, Rhys and Ian Dew-Becker, "Long-run risk is the worst-case scenario," *American Economic Review*, 2016, *106* (9), 2494–2527.

Bondt, Werner PM De, "Betting on trends: Intuitive forecasts of financial risk and return," *International Journal of Forecasting*, 1993, *9* (3), 355–371.

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

— **,** — **,** — **, and** — , "Belief overreaction and stock market puzzles," *Journal of Political Economy*, 2024, *132* (5), 1450–1484.

— **,** — **, Rafael La Porta, and Andrei Shleifer**, "Finance Without (Exotic) Risk," 2024.

Borovička, Jaroslav, "Survival and long-run dynamics with heterogeneous beliefs under recursive preferences," *Journal of Political Economy*, 2020, *128* (1), 206–251.

— **, Lars Peter Hansen, and José A Scheinkman**, "Misspecified recovery," *The Journal of Finance*, 2016, *71* (6), 2493–2544.

Bretscher, Lorenzo, Aytek Malkhozov, Andrea Tamoni, and Haoxi Yang, "Distorted Beliefs and Asset Prices," *Available at SSRN 4947547*, 2024.

Buraschi, Andrea and Alexei Jiltsov, "Model uncertainty and option markets with heterogeneous beliefs," *The Journal of Finance*, 2006, *61* (6), 2841–2897.

Cassella, Stefano, Te-Feng Chen, Huseyin Gulen, and Yan Liu, "Extracting extrapolative beliefs from market prices: An augmented present-value approach," *Journal of Financial Economics*, 2025, *164*, 103986.

Chabakauri, Georgy, "Asset pricing with heterogeneous preferences, beliefs, and portfolio constraints," *Journal of Monetary Economics*, 2015, *75*, 21–34.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, "Measuring "Dark Matter" in Asset Pricing Models," *The Journal of Finance*, 2024, *79* (2), 843–902.

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

Chen, Luyang, Markus Pelger, and Jason Zhu, "Deep learning in asset pricing," *Management Science*, 2024, *70* (2), 714–750.

Chen, Xiaohong, Lars Peter Hansen, and Peter G Hansen, "Robust inference for moment condition models without rational expectations," *Journal of Econometrics*, 2024, p. 105653.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A Lochstoer, "Asset pricing when 'this time is different'," *The Review of Financial Studies*, 2017, *30* (2), 505–535.

Couts, Spencer J, Andrei S Gonçalves, Yicheng Liu, and Johnathan Loudis, "Institutional Investors' Subjective Risk Premia: Time Variation and Disagreement," 2024.

Dahlquist, Magnus and Markus Ibert, "Equity return expectations and portfolios: Evidence from large asset managers," *The Review of Financial Studies*, 2024, *37* (6), 1887–1928.

David, Alexander, "Heterogeneous beliefs, speculation, and the equity premium," *The Journal of Finance*, 2008, *63* (1), 41–83.

Detemple, Jerome and Shashidhar Murthy, "Intertemporal asset pricing with heterogeneous beliefs," *Journal of Economic Theory*, 1994, *62* (2), 294–320.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal, "Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility," *The Journal of Finance*, 2009, *64* (2), 579–629.

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen, "Asset prices and portfolio choice with learning from experience," *The Review of Economic Studies*, 2018, *85* (3), 1752–1780.

Epstein, Larry G and Stanley E Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 1989, pp. 937–969.

Fuster, Andreas, David Laibson, and Brock Mendel, "Natural expectations and macroeconomic fluctuations," *Journal of Economic Perspectives*, 2010, *24* (4), 67–84.

Gârleanu, Nicolae and Stavros Panageas, "Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing," *Journal of Political Economy*, 2015, *123* (3), 670–685.

Ghosh, Anisha and Guillaume Roussellet, "Identifying beliefs from asset prices," 2023. Unpublished working paper.

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

Gürkaynak, Refet S, Brian Sack, and Jonathan H Wright, "The US Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 2007, *54* (8), 2291–2304.

Hansen, Lars Peter and José A Scheinkman, "Long-term risk: An operator approach," *Econometrica*, 2009, *77* (1), 177–234.

— **and Ravi Jagannathan**, "Implications of security market data for models of dynamic economies," *Journal of Political Economy*, 1991, *99* (2), 225–262.

Harrison, J Michael and David M Kreps, "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, 1979, *20* (3), 381–408.

Jensen, Theis Ingerslev, Bryan T Kelly, Semyon Malamud, and Lasse Heje Pedersen, "Machine learning and the implementable efficient frontier," *Swiss Finance Institute Research Paper*, 2024, (22-63).

Kasa, Kenneth, Todd B Walker, and Charles H Whiteman, "Heterogeneous beliefs and tests of present value models," *Review of Economic Studies*, 2014, *81* (3), 1137–1163.

Kazemi, Hossein B, "An intemporal model of asset prices in a Markov economy with a limiting stationary distribution," *The Review of Financial Studies*, 1992, *5* (1), 85–104.

Khorrami, Paymon, "Entry and slow-moving capital: using asset markets to infer the costs of risk concentration," 2023. Unpublished working paper.

Kim, Don H and Athanasios Orphanides, "Term structure estimation with survey data on interest rate forecasts," *Journal of Financial and Quantitative Analysis*, 2012, *47* (1), 241–272.

Kogan, Leonid, Stephen A Ross, Jiang Wang, and Mark M Westerfield, "The price impact and survival of irrational traders," *The Journal of Finance*, 2006, *61* (1), 195–229.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, "Interpreting factor models," *The Journal of Finance*, 2018, *73* (3), 1183–1223.

Manela, Asaf and Alan Moreira, "News implied volatility and disaster concerns," *Journal of Financial Economics*, 2017, *123* (1), 137–162.

Martin, Ian, "What is the Expected Return on the Market?," *The Quarterly Journal of Economics*, 2017, *132* (1), 367–433.

— **and Dimitris Papadimitriou**, "Sentiment and speculation in a market with heterogeneous beliefs," *American Economic Review*, 2022, *112* (8), 2465–2517.

— **and Steve Ross**, "The long bond," 2013. London School of Economics and MIT Sloan Mimeo.

Moreira, Alan and Tyler Muir, "Volatility-managed portfolios," *The Journal of Finance*, 2017, *72* (4), 1611–1644.

Nagel, Stefan and Zhengyang Xu, "Asset pricing with fading memory," *The Review of Financial Studies*, 2022, *35* (5), 2190–2245.

— **and** — , "Dynamics of subjective risk premia," *Journal of Financial Economics*, 2023, *150* (2), 103713.

Newey, Whitney K and Kenneth D West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica: Journal of the Econometric Society*, 1987, pp. 703–708.

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

_ and _ , "Which Subjective Expectations Explain Asset Prices?," *The Review of Financial Studies*, 2024, *37* (6), 1929–1978.

_ , Xiao Han, and Sean Myers, "The Return of Return Dominance: Decomposing the Cross-section of Prices," 2023.

Panageas, Stavros, "The implications of heterogeneity and inequality for asset pricing," *Foundations and Trends in Finance*, 2020, *12* (3), 199–275.

Qin, Likuan and Vadim Linetsky, "Positive eigenfunctions of Markovian pricing operators: Hansen-Scheinkman factorization, Ross recovery, and long-term pricing," *Operations Research*, 2016, *64* (1), 99–117.

_ and _ , "Long-term risk: A martingale approach," *Econometrica*, 2017, *85* (1), 299–312.

Roll, Richard, "Evidence on the "growth-optimum" model," *The Journal of Finance*, 1973, *28* (3), 551–566.

Ross, Steve, "The recovery theorem," *The Journal of Finance*, 2015, *70* (2), 615–648.

Scheinkman, Jose A and Wei Xiong, "Overconfidence and speculative bubbles," *Journal of Political Economy*, 2003, *111* (6), 1183–1220.

Yan, Hongjun, "Natural selection in financial markets: Does it work?," *Management Science*, 2008, *54* (11), 1935–1950.

Zapatero, Fernando, "Effects of financial innovations on market volatility when beliefs are heterogeneous," *Journal of Economic Dynamics and Control*, 1998, *22* (4), 597–626.

Appendix

A Proofs

A.1 Proof of Lemma 1

We prove only the subjective versions; the objective bounds follow from those by taking $B \equiv 1$ (no belief distortion). Note by the pricing equation (5), the permanent-transitory factorization (8), the fact $\tilde{G} = G$, and the long bond result (10) that

$$\begin{aligned} 0 \leq \tilde{L}_t\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t} R_{t+1}\Big) &= -\tilde{\mathbb{E}}_t \log\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}\Big) - \tilde{\mathbb{E}}_t \log R_{t+1} \\ &= -\tilde{\mathbb{E}}_t \log\Big(\frac{G_{t+1}}{G_t}\Big) - \tilde{\mathbb{E}}_t \log\Big(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}\Big) - \tilde{\mathbb{E}}_t \log R_{t+1} \\ &= \tilde{\mathbb{E}}_t \log R^{\infty}_{t+1} + \tilde{L}_t\Big(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}\Big) - \tilde{\mathbb{E}}_t \log R_{t+1} \end{aligned}$$

Rearranging leads to (16). To derive (14), start with the identity

$$\tilde{L}_t\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}\Big) = \tilde{L}_t\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t} R^f_t\Big) = \tilde{\mathbb{E}}_t[\log R^{\infty}_{t+1}] - \log R^f_t + \tilde{L}_t\Big(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}\Big) \tag{A.1}$$

and then use (16).

A.2 Proof of Proposition 1

To prove (20), use the UEI property to get $L(\frac{\tilde{H}_{t+1}}{\tilde{H}_t}) \geq \mathbb{E}[\tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})]$. Combining this result with the bound in (16), we obtain (20).

The same argument with $\frac{\tilde{S}_{t+1}}{\tilde{S}_t} R^f_t$ in place of $\frac{\tilde{H}_{t+1}}{\tilde{H}_t}$ leads to $L(\frac{\tilde{S}_{t+1}}{\tilde{S}_t} R^f_t) \geq \mathbb{E}[\tilde{L}_t(\frac{\tilde{S}_{t+1}}{\tilde{S}_t} R^f_t)]$. Because R^f_t is conditionally risk-free, we have $\tilde{L}_t(\frac{\tilde{S}_{t+1}}{\tilde{S}_t} R^f_t) = \tilde{L}_t(\frac{\tilde{S}_{t+1}}{\tilde{S}_t})$. Combining these results with the bound in (14), we obtain (19).

Finally, we can deduce (17)-(18) from the subjective versions by considering $B \equiv 1$ (i.e., no belief distortion), so that all objects and expectations with tildes become objective versions, and then applying the law of iterated expectations.

A.3 Proof of Proposition 2

First, make note of the following key result: the inequalities in (13) and (15) become equalities with $R_{t+1} = R^*_{t+1}$. Similarly, the subjective versions in (14) and (16) become equalities with $R_{t+1} = \tilde{R}^*_{t+1}$.

Now, let us characterize the "wedges" between the bounds in Proposition 1: For the objective bounds, we obtain

$$\begin{aligned} \omega(R) &:= L\Big(\frac{S_{t+1}}{S_t} R^f_t\Big) - \mathbb{E}[\log R_{t+1} - \log R^f_t] \\ &= \mathbb{E}[\log R^*_{t+1} - \log R^f_t] - \mathbb{E}[\log R_{t+1} - \log R^f_t] = \mathbb{E}[\delta_t(R)] \end{aligned}$$

For the subjective bounds, we also use the UEI assumption to obtain

$$\begin{aligned}
\tilde{\omega}(R) &:= L\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}R_t^f\Big) - \mathbb{E}\big[\tilde{\mathbb{E}}_t[\log R_{t+1} - \log R_t^f]\big] \\
&\geq \mathbb{E}\Big[\tilde{L}_t\Big(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}R_t^f\Big)\Big] - \mathbb{E}\big[\tilde{\mathbb{E}}_t[\log R_{t+1} - \log R_t^f]\big] \\
&= \mathbb{E}\big[\tilde{\mathbb{E}}_t[\log R_{t+1}^* - \log R_t^f]\big] - \mathbb{E}\big[\tilde{\mathbb{E}}_t[\log R_{t+1} - \log R_t^f]\big] = \mathbb{E}[\tilde{\delta}_t(R)]
\end{aligned}$$

Given two returns $R, \tilde{R}$, we use the assumption that R is closer to objective growth-optimal than $\tilde{R}$ is to subjective growth-optimal, i.e., $\mathbb{E}[\delta_t(R)] \leq \mathbb{E}[\tilde{\delta}_t(\tilde{R})]$, to obtain

$$\omega(R) \leq \tilde{\omega}(\tilde{R}).$$

Therefore, compute

$$\frac{L(\frac{\tilde{S}_{t+1}}{\tilde{S}_t}R_t^f)}{L(\frac{S_{t+1}}{S_t}R_t^f)} = \frac{\mathbb{E}\big[\tilde{\mathbb{E}}_t[\log \tilde{R}_{t+1} - \log R_t^f]\big] + \tilde{\omega}(\tilde{R})}{\mathbb{E}[\log R_{t+1} - \log R_t^f] + \omega(R)} \geq \frac{\mathbb{E}\big[\tilde{\mathbb{E}}_t[\log \tilde{R}_{t+1} - \log R_t^f]\big] + \omega(R)}{\mathbb{E}[\log R_{t+1} - \log R_t^f] + \omega(R)}$$

Note that the latter ratio is increasing (decreasing) in $\omega(R)$ if the ratio is smaller (larger) than one. This leads to the result in (21). To obtain (22), we repeat a very similar argument to obtain

$$\frac{L(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{L(\frac{H_{t+1}}{H_t})} \geq \frac{\mathbb{E}\big[\tilde{\mathbb{E}}_t[\log \tilde{R}_{t+1} - \log R_t^\infty]\big] + \omega_H(R)}{\mathbb{E}[\log R_{t+1} - \log R_t^\infty] + \omega_H(R)},$$

where $\omega_H(R) := L\big(\frac{H_{t+1}}{H_t}\big) - \mathbb{E}[\log R_{t+1} - \log R_t^\infty]$ is the corresponding wedge for the permanent component.

A.4 Proof of Corollary 1

See Alvarez and Jermann (2005), Proposition 2.

A.5 Proof of Corollary 2

Throughout, suppose R is a return such that $\tilde{\mathbb{E}}_t \log R_{t+1} > \log R_t^f$. From (A.1), we have

$$\frac{\tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{\tilde{L}_t(\frac{\tilde{S}_{t+1}}{\tilde{S}_t})} = \frac{\tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{\tilde{\mathbb{E}}_t \log R_{t+1}^\infty - \log R_t^f + \tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}$$

Note that this ratio is increasing (decreasing) in $\tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})$ if it is smaller (larger) than one. So if $\tilde{\mathbb{E}}_t \log R_{t+1}^\infty > \log R_t^f$, we use (16) to get

$$1 \geq \frac{\tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{\tilde{L}_t(\frac{\tilde{S}_{t+1}}{\tilde{S}_t})} \geq \frac{\tilde{\mathbb{E}}_t[\log R_{t+1} - \log R_{t+1}^\infty]}{\tilde{\mathbb{E}}_t \log R_{t+1} - \log R_t^f}$$

On the other hand, if $\tilde{\mathbb{E}}_t \log R_{t+1}^\infty < \log R_t^f$, we use (16) to get

$$1 \leq \frac{\tilde{L}_t(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{\tilde{L}_t(\frac{\tilde{S}_{t+1}}{\tilde{S}_t})} \leq \frac{\tilde{\mathbb{E}}_t[\log R_{t+1} - \log R_{t+1}^\infty]}{\tilde{\mathbb{E}}_t \log R_{t+1} - \log R_t^f}$$

Combine these results with the non-negativity of entropies to obtain the result.

A.6 UEI holds in continuous-time Brownian environments

The Unconditional Entropy Inequality (UEI) assumption says that all higher moments beyond the mean are perceived equally by the subjective and objective probability. In this section, I show, as claimed by Footnote 3, that the UEI automatically holds in continuous-time Brownian environments. I model the dynamics of a variable $\log X_t$ via an Itô process and prove in such case that

$$\lim_{\Delta \to 0} \Delta^{-1} \tilde{L}_t(X_{t+\Delta}) = \lim_{\Delta \to 0} \Delta^{-1} L_t(X_{t+\Delta}).$$

i.e., local versions of the entropy operator coincide.

Suppose

$$d \log X_t = \alpha_t dt + \nu_t \cdot dW_t,$$

where W is a Brownian motion under $\mathbb{P}$. By Girsanov's theorem, the dynamics under the subjective measure are

$$d \log X_t = \tilde{\alpha}_t dt + \nu_t \cdot d\tilde{W}_t,$$

where $\tilde{W}$ is a Brownian motion under $\tilde{\mathbb{P}}$ and $\tilde{\alpha}_t$ is another adapted process. Compute

$$L_t(X_{t+\Delta}) = \log(\mathbb{E}_t[\exp(\int_t^{t+\Delta} (\alpha_s + \frac{1}{2}|\nu_s|^2) ds)]) - \int_t^{t+\Delta} \alpha_s ds$$
$$\tilde{L}_t(X_{t+\Delta}) = \log(\tilde{\mathbb{E}}_t[\exp(\int_t^{t+\Delta} (\tilde{\alpha}_s + \frac{1}{2}|\nu_s|^2) ds)]) - \int_t^{t+\Delta} \tilde{\alpha}_s ds,$$

where $\mathbb{E}$ and $\tilde{\mathbb{E}}$ are alternative expectation operators. Dividing both metrics by Δ and taking $\Delta \to 0$, and using L'Hôpital's rule, yields

$$\lim_{\Delta \to 0} \Delta^{-1} L_t(X_{t+\Delta}) = \alpha_t + \frac{1}{2}|\nu_t|^2 - \alpha_t = \frac{1}{2}|\nu_t|^2 = \tilde{\alpha}_t + \frac{1}{2}|\nu_t|^2 - \tilde{\alpha}_t = \lim_{\Delta \to 0} \Delta^{-1} \tilde{L}_t(X_{t+\Delta})$$

Thus, the two metrics coincide locally.

A.7 Proof of Proposition 3

Using the calculations in Appendix B, using the notation $\Delta X = \frac{X_{t+1}}{X_t}$ for any variable X, and using the definition of the perceived persistence $\tilde{\rho} = 1 - (1 - \phi)\lambda$, we have the following connection between the two volatility ratios

$$\frac{L(\Delta G)}{L(\Delta S)} = K \times \frac{L(\Delta B)}{L(\Delta S)}, \quad \text{where} \quad K := \frac{2(1 - \tilde{\rho} + \phi \nu_1 / \sigma_c)}{(1 - \tilde{\rho})^2 \sigma_c^2}$$

This automatically proves the result that $\frac{L(\Delta G)}{L(\Delta S)}$ and $\frac{L(\Delta B)}{L(\Delta S)}$ move in the same direction in response to any parameter not explicitly included in K. Thus, the result is proved for γ, β, ν_2.

For the remaining parameter ϕ, we proceed directly as follows. First, from the fact that K is increasing in ϕ, holding $\tilde{\rho}$ fixed, it suffices to show that $\frac{L(\Delta B)}{L(\Delta S)}$ is increasing in ϕ, because that will imply that $\frac{L(\Delta G)}{L(\Delta S)}$ is also increasing in ϕ. Second, using the expression in Appendix B and noting $\text{Var}[r_t] = \text{Var}[x_t]$, we may write

$$\frac{L(\Delta B)}{L(\Delta S)} = \frac{1/\sigma_c^2}{\|\tilde{\pi}\|^2 / \text{Var}[x_t] + 1 + 1/\sigma_c^2}.$$

Consequently, $\frac{L(\Delta B)}{L(\Delta S)}$ is increasing in ϕ if and only if $\|\tilde{\pi}\|^2/\mathrm{Var}[x_t]$ is decreasing in ϕ. Using the expressions $\mathrm{Var}[x_t] = \frac{\phi^2\|\nu\|^2}{1-(\tilde{\rho}-\phi\nu_1/\sigma_c)^2}$ and $\tilde{\pi} = \gamma\sigma_c\binom{1}{0} + \frac{(\gamma-1)\beta}{1-\beta\tilde{\rho}}\phi\nu$, this ratio is

$$\frac{\|\tilde{\pi}\|^2}{\mathrm{Var}[x_t]} = \Big[1 - \big(\tilde{\rho} - \frac{\phi\nu_1}{\sigma_c}\big)^2\Big] \frac{[\gamma\sigma_c + \frac{(\gamma-1)\beta\phi\nu_1}{1-\beta\tilde{\rho}}]^2 + [\frac{(\gamma-1)\beta\phi\nu_2}{1-\beta\tilde{\rho}}]^2}{\phi^2\|\nu\|^2}$$

Differentiate $\|\tilde{\pi}\|^2/\mathrm{Var}[x_t]$ with respect to ϕ, holding $\tilde{\rho}$ fixed, to get

$$\frac{\partial}{\partial\phi}\frac{\|\tilde{\pi}\|^2}{\mathrm{Var}[x_t]}\Big|_{\tilde{\rho}\text{ fixed}} = 2\frac{1}{\phi^3\|\nu\|^2}\Big\{\Big(\tilde{\rho} - \frac{\phi\nu_1}{\sigma_c}\Big)\frac{\phi\nu_1}{\sigma_c}\|\tilde{\pi}\|^2 - \Big[1 - \big(\tilde{\rho} - \frac{\phi\nu_1}{\sigma_c}\big)^2\Big]\Big(\gamma\sigma_c + \frac{(\gamma-1)\beta\phi\nu_1}{1-\beta\tilde{\rho}}\Big)\gamma\sigma_c\Big\}$$

This expression is clearly negative for ϕ small enough or for $\tilde{\rho}$ small enough.

A.8 Proof of Proposition 4

See Appendix C for the expressions for the conditional entropies to obtain (51).

B Derivations for the models in Section 4

Value function in baseline long-run risk model of Section 4.1. Guess that the value function is given by the form $U_t = \log C_t + u_0 + u_x x_t$, and obtain the values of u_0 and u_x by substituting this guess into the Bellman equation (28) and using the method of undetermined coefficients. We report $u_x = \frac{\beta}{1-\beta\rho}$ in the main text. The "constant" term u_0 (not reported in the main text) is

$$u_0 = \frac{\beta}{1-\beta}\Big[(1-\rho)\bar{x}u_x + \frac{1}{2}(1-\gamma)\big([\sigma_c + (1-\kappa^2)^{1/2}\sigma_x u_x]^2 + \kappa^2\sigma_x^2 u_x^2\big)\Big]$$

Permanent-transitory decomposition in long-run risk model of Section 4.1. Guess that the transitory piece is given by $G_t = e^{\eta t - \zeta(x_t - x_0)}$ for some constants (η, ζ), and then substitute the guess $S_t = G_t H_t$ into the SDF formula. This yields

$$\begin{aligned}
\frac{H_{t+1}}{H_t} &= \exp(-\eta + \zeta(x_{t+1} - x_t))\frac{S_{t+1}}{S_t} \\
&= \exp\Big\{ -\eta - r_t - \big(\pi - \zeta\sigma_x\big(\tfrac{(1-\kappa^2)^{1/2}}{\kappa}\big)\big)\cdot W_{t+1} - \frac{1}{2}\|\pi\|^2 + \zeta(1-\rho)(\bar{x} - x_t)\Big\}
\end{aligned}$$

Using the conjecture that H is a martingale allows us to solve for η and ζ with the method of undetermined coefficients. These values are then

$$\begin{aligned}
&\zeta = -(1-\rho)^{-1} \\
&\eta = \log(\beta) - \frac{1}{2}\sigma_c^2 + \gamma\sigma_c^2 + \frac{(\gamma-1)\beta}{1-\beta\rho}(1-\kappa^2)^{1/2}\sigma_x\sigma_c + \zeta(1-\rho)\bar{x} + \frac{1}{2}\zeta^2\sigma_x^2 - \zeta\sigma_x\big(\tfrac{(1-\kappa^2)^{1/2}}{\kappa}\big)\cdot\pi.
\end{aligned}$$

Volatilities in long-run risk model of Section 4.1. The annual average SDF volatility is

$$\mathbb{E}\Big[2L_t\Big(\frac{S_{t+1}}{S_t}\Big)\Big]^{-1/2} = \|\pi\|$$

The sizes of the permanent and transitory components are given by

$$\begin{aligned}
\frac{L(\frac{H_{t+1}}{H_t})}{L(\frac{S_{t+1}}{S_t})} &= \frac{\|\pi + \frac{\sigma_x}{1-\rho}\big(\tfrac{(1-\kappa^2)^{1/2}}{\kappa}\big)\|^2}{\|\pi\|^2 + \mathrm{Var}[r_t]} \\
\frac{L(\frac{G_{t+1}}{G_t})}{L(\frac{S_{t+1}}{S_t})} &= \frac{\frac{2}{1-\rho}\mathrm{Var}[r_t]}{\|\pi\|^2 + \mathrm{Var}[r_t]}.
\end{aligned}$$

For all of these expressions, note that $\mathrm{Var}[r_t] = \mathrm{Var}[x_t]$, i.e., the objective variance of agent's growth perceptions. Theoretically, one can show that that $L(\Delta H)/L(\Delta S)$ is always greater than 1 if $\gamma \geq 1$.[20]

Volatilities in the growth-extrapolation model of Section 4.2. The annual average SDF volatility is

$$\mathbb{E}\Big[2L_t\Big(\frac{S_{t+1}}{S_t}\Big)\Big]^{1/2} = \Big(\mathbb{E}\|\tilde{\pi} - \mu_t\|^2\Big)^{1/2} = \Big(\|\tilde{\pi}\|^2 + \frac{\mathrm{Var}[r_t]}{\sigma_c^2}\Big)^{1/2}.$$

[20] The unconditional variance of r_t is computed as $\sigma_x^2/(1-\rho^2)$ using its AR(1) dynamics inherited from x_t. Then, using the fact that $\pi \geq 0$ when $\gamma \geq 1$, followed by the assumption $0 < \rho < 1$, we have $\|\pi + \frac{\sigma_x}{1-\rho}\big(\tfrac{(1-\kappa^2)^{1/2}}{\kappa}\big)\|^2 \geq \|\pi\|^2 + (\frac{\sigma_x}{1-\rho})^2 \geq \|\pi\|^2 + \frac{\sigma_x^2}{1-\rho^2} = \|\pi\|^2 + \mathrm{Var}[r_t]$.

The sizes of the permanent and transitory components are given by

$$\frac{L(\frac{H_{t+1}}{H_t})}{L(\frac{S_{t+1}}{S_t})} = \frac{\|\tilde{\pi} + \frac{\phi\nu}{1-\tilde{\rho}}\|^2 + \frac{1}{\sigma_c^2}\text{Var}[r_t]}{\|\tilde{\pi}\|^2 + (1 + \frac{1}{\sigma_c^2})\text{Var}[r_t]}$$

$$\frac{L(\frac{G_{t+1}}{G_t})}{L(\frac{S_{t+1}}{S_t})} = \frac{\frac{2(1-\tilde{\rho}+\frac{\phi\nu_1}{\sigma_c})}{(1-\tilde{\rho})^2}\text{Var}[r_t]}{\|\tilde{\pi}\|^2 + (1 + \frac{1}{\sigma_c^2})\text{Var}[r_t]}$$

Finally, the ratio of subjective-to-objective SDF volatility is given by

$$\frac{L(\frac{\tilde{S}_{t+1}}{\tilde{S}_t})}{L(\frac{S_{t+1}}{S_t})} = \frac{\|\tilde{\pi}\|^2 + (1 - \frac{\tilde{\pi}_1}{\sigma_c})^2\text{Var}[r_t]}{\|\tilde{\pi}\|^2 + (1 + \frac{1}{\sigma_c^2})\text{Var}[r_t]}$$

and its permanent counterpart by

$$\frac{L(\frac{\tilde{H}_{t+1}}{\tilde{H}_t})}{L(\frac{H_{t+1}}{H_t})} = \frac{\|\tilde{\pi} + \frac{\phi\nu}{1-\tilde{\rho}}\|^2 + (\tilde{\pi}_1 + \frac{\phi\nu_1}{1-\tilde{\rho}})^2\frac{1}{\sigma_c^2}\text{Var}[r_t]}{\|\tilde{\pi} + \frac{\phi\nu}{1-\tilde{\rho}}\|^2 + \frac{1}{\sigma_c^2}\text{Var}[r_t]}$$

Finally, the size of belief volatility in the SDF is given by

$$\frac{L(\frac{B_{t+1}}{B_t})}{L(\frac{S_{t+1}}{S_t})} = \frac{\frac{1}{\sigma_c^2}\text{Var}[r_t]}{\|\tilde{\pi}\|^2 + (1 + \frac{1}{\sigma_c^2})\text{Var}[r_t]}$$

To compute all these volatility ratios, we can calculate them directly or use Lemma B.1 below with $M \in \{\tilde{S}, S, \tilde{H}, H, G, B\}$. For all of these expressions, note that $\text{Var}[r_t] = \text{Var}[x_t]$. Furthermore, using $\tilde{W}_{t+1} = W_{t+1} - \mu_t$, we can show that x_t also follows an AR(1) under the objective measure, but with modified persistence:

$$x_{t+1} - \bar{x} = \Big(1 - \lambda(1-\phi) - \frac{\phi\nu_1}{\sigma_c}\Big)(x_t - \bar{x}) + \phi\nu \cdot W_{t+1},$$

and so its unconditional moments are $\mathbb{E}[x_t] = \bar{x}$ and $\text{Var}[x_t] = \frac{\phi^2\|\nu\|^2}{1-(1-\lambda(1-\phi)-\frac{\phi\nu_1}{\sigma_c})^2}$.

Solution to the CRRA growth-extrapolation model of Section 4.3. Following the same procedures as above, we can compute marginal utility, its perceived permanent component, the objective SDF, the objective permanent component, and the transitory component as

$$\begin{aligned}
\frac{\tilde{S}_{t+1}}{\tilde{S}_t} &= \exp\Big\{ - r_t - \frac{1}{2}\|\tilde{\pi}\|^2 - \tilde{\pi}\cdot\tilde{W}_{t+1}\Big\} \\
\frac{\tilde{H}_{t+1}}{\tilde{H}_t} &= \exp\Big\{ - \frac{1}{2}\Big\|\tilde{\pi} + \frac{\gamma\phi\nu}{1-\tilde{\rho}}\Big\|^2 - \Big(\tilde{\pi} + \frac{\gamma\phi\nu}{1-\tilde{\rho}}\Big)\cdot\tilde{W}_{t+1}\Big\} \\
\frac{S_{t+1}}{S_t} &= \exp\Big\{ - r_t - \frac{1}{2}\|\tilde{\pi} - \mu_t\|^2 - (\tilde{\pi} - \mu_t)\cdot W_{t+1}\Big\} \\
\frac{H_{t+1}}{H_t} &= \exp\Big\{ - \frac{1}{2}\Big\|\tilde{\pi} - \mu_t + \frac{\gamma\phi\nu}{1-\tilde{\rho}}\Big\|^2 - \Big(\tilde{\pi} - \mu_t + \frac{\gamma\phi\nu}{1-\tilde{\rho}}\Big)\cdot W_{t+1}\Big\} \\
\frac{G_{t+1}}{G_t} &= \exp\Big[\tilde{\eta} + \frac{\gamma}{1-\tilde{\rho}}(x_{t+1} - x_t)\Big],
\end{aligned}$$

for some constant $\tilde{\eta}$. Note that the belief distortion B is identical in this CRRA model as the Epstein-Zin version above. To compute volatility ratios, we again use Lemma B.1 for a generic $M \in \{\tilde{S}, S, \tilde{H}, H, G, B\}$.

The volatility ratios in these models are all easily computable analytically, because of the log-normality of the equilibrium solution. We have the following tool for this purpose. Every object of interest in this model takes the form of some M in (B.1) below. Therefore, applying this lemma to the specific forms of S, H, $\tilde{S}$, $\tilde{H}$, G, and B, we can obtain any volatility ratio we like. Note that to apply this to objects like $\tilde{S}$ and $\tilde{H}$, which are written in the text as being driven by the subjective shocks $\tilde{W}$, we first need to write them in terms of the objective shocks by using the identity $\tilde{W}_{t+1} = W_{t+1} - \mu_t$.

Lemma B.1. *For some scalars a, b, q and vector p, consider the process*

$$\frac{M_{t+1}}{M_t} = \exp\left\{a + bx_t - [p + q\mu_t] \cdot W_{t+1} - \frac{1}{2}\|p + q\mu_t\|^2\right\}, \tag{B.1}$$

where x_t follows an AR(1) with long-run mean $\bar{x}$, and where $\mu_t = \frac{x_t - \bar{x}}{\sigma_c}$. Then,

$$L\left(\frac{M_{t+1}}{M_t}\right) = \log \mathbb{E}\left[e^{a+bx_t}\right] - (a + b\bar{x}) + \frac{1}{2}\mathbb{E}\|p + q\mu_t\|^2 = \frac{1}{2}\left(b^2 + \frac{q^2}{\sigma_c^2}\right) Var[x_t] + \frac{1}{2}\|p\|^2. \tag{B.2}$$

Proof. This follows from direct calculation of $L(\Delta M) = \log \mathbb{E}[\Delta M] - \mathbb{E}[\log \Delta M]$ using the fact that $\mathbb{E}[\mu_t] = 0$, $\text{Var}[\binom{1}{0} \cdot \mu_t] = \frac{1}{\sigma_c^2}\text{Var}[x_t]$, and $\text{Var}[x_t]$ is the stationary variance of the AR(1) x_t. $\square$

C Derivations for the models in Section 5

Equilibrium derivation. First, we solve the model up to a coupled pair of ODEs.

Proposition C.1. *Assume a solution exists to the system of ODEs (C.6) below. Then, there exists a bubble-free equilibrium that is Markovian in the pessimists' consumption share x_t. The equilibrium market price of risk and interest rate are given by*

$$\pi_t = \gamma\sigma_c\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) + x_t\beta\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right)$$
$$r_t = \rho + \gamma\delta\Big(1 - \sum_\theta \eta^\theta \frac{c^\theta_{t,t}}{C_t}\Big) + \gamma(\mu_c - \sigma_c^2) - \frac{1}{2}\frac{1+\gamma}{\gamma}\|\pi_t\|^2 + \sigma_c\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\cdot\pi_t$$
$$- x_t\frac{\beta}{\gamma}\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right)\cdot\Big[\gamma\sigma_c\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) + \frac{1}{2}(1-\gamma)\beta\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right) - \pi_t\Big].$$

The drift and diffusion of pessimist's consumption share are given by

$$\mu_{x,t} = \delta\Big[(1-x_t)\frac{\eta^1 c^1_{t,t}}{C_t} - x_t\frac{\eta^0 c^0_{t,t}}{C_t}\Big] + x_t(1-x_t)\frac{\beta}{\gamma^2}\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right)\cdot\Big[\gamma\sigma_c\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) + \frac{1}{2}(1-\gamma)\beta\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right) - \pi_t\Big]$$
$$\sigma_{x,t} = -\frac{1}{\gamma}x_t(1-x_t)\beta\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right)$$

The expressions for the newborn consumption shares $c^\theta_{t,t}/C_t$ are given in the appendix.

Proof of Proposition C.1. The FOC for a type-θ agent born at time u is

$$\frac{c^\theta_{t,u}}{c^\theta_{u,u}} = e^{-\frac{\rho}{\gamma}(t-u)}\Big(\frac{B^\theta_t}{B^\theta_u}\Big)^{\frac{1}{\gamma}}\Big(\frac{S_t}{S_u}\Big)^{-\frac{1}{\gamma}} \tag{C.1}$$

Plug this into the definition of aggregate type-θ consumption $c^\theta_t := \eta^\theta \int_{-\infty}^t \delta e^{-\delta(t-u)} c^\theta_{t,u} du$ to obtain the following expression for the consumption share $x^\theta_t := c^\theta_t/C_t$:

$$x^\theta_t = \frac{\eta^\theta \int_{-\infty}^t \delta e^{-(\delta+\frac{\rho}{\gamma})(t-u)} c^\theta_{u,u}\big(\frac{B^\theta_t}{B^\theta_u}\big)^{\frac{1}{\gamma}}\big(\frac{S_t}{S_u}\big)^{-\frac{1}{\gamma}} du}{C_t} \tag{C.2}$$

Applying Itô's formula to this expression, we obtain

$$dx^\theta_t = \delta\frac{\eta^\theta c^\theta_{t,t}}{C_t}dt - \Big(\delta + \frac{\rho}{\gamma}\Big)x^\theta_t dt + \frac{1}{\gamma}x^\theta_t\Big(\frac{dB^\theta_t}{B^\theta_t} - \frac{dS_t}{S_t}\Big) \tag{C.3}$$
$$+ x^\theta_t\Big[\frac{1}{2\gamma^2}\Big((1-\gamma)\frac{d[B^\theta]_t}{(B^\theta_t)^2} + (1+\gamma)\frac{d[S]_t}{S_t^2} - 2\frac{d[B^\theta,S]_t}{B^\theta_t S_t}\Big) - \frac{dC_t}{C_t} + \frac{d[C]_t}{C_t^2} - \frac{1}{\gamma}\Big(\frac{d[B^\theta,C]_t}{B^\theta_t C_t} - \frac{d[S,C]_t}{S_t C_t}\Big)\Big]$$

Substituting $dB^\theta_t = B^\theta_t\theta\beta\left(\begin{smallmatrix}1\\\chi\end{smallmatrix}\right)\cdot dW_t$, $dS_t = -S_t[r_t dt + \pi_t\cdot dW_t]$, and $dC_t = C_t[\mu_c dt + \sigma_c\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\cdot dW_t]$, and then using the market clearing condition $\sum dx^\theta_t = 0$, we obtain the equations for π_t and r_t given in the proposition. Substituting those results back into (C.3) for $\theta = 1$, we obtain the expressions for $\mu_{x,t}$ and $\sigma_{x,t}$ as well.

Next, substitute the FOC (C.1) into the lifetime budget constraint (46) to obtain

$$\frac{c^\theta_{t,t}}{C_t} = \frac{\kappa p_t}{\delta q_{\theta,t}}, \tag{C.4}$$

where we have defined

$$
\begin{aligned}
p_t &:= \mathbb{E}_t\Big[\int_t^\infty e^{-\kappa(u-t)}\frac{S_u}{S_t}\frac{C_u}{C_t}du\Big]\\
q_{\theta,t} &:= \mathbb{E}\Big[\int_t^\infty e^{-(\delta+\frac{\rho}{\gamma})(u-t)}\big(\frac{B_u^\theta}{B_t^\theta}\big)^{\frac{1}{\gamma}}\big(\frac{S_u}{S_t}\big)^{1-\frac{1}{\gamma}}du\Big]
\end{aligned}
$$

As these are present values, we are already imposing a conjecture that the equilibrium is bubble-free, as stated in the proposition. Note that $p_t = P_t/C_t$ is the aggregate price-dividend ratio in the economy, if the equilibrium is bubble-free. Given the homogeneity of preferences and budget sets, all alive type-θ agents consume the same fraction of their wealth. Hence, the type-θ wealth-to-consumption ratio n_t^θ/c_t^θ is equal to the wealth-to-consumption ratio of a type-θ newborn. A newborn's initial wealth is $P_{t,t}/\delta = \kappa p_t C_t/\delta$ and their consumption is $c_{t,t}^\theta$, which by (C.4) implies that the type-θ wealth-consumption ratio is $q_{\theta,t}$. With this result, we use asset market clearing $\sum_\theta n_t^\theta = P_t$ to obtain that

$$
\sum_\theta x_t^\theta q_{\theta,t} = p_t. \tag{C.5}
$$

Since $x_t^0 = 1 - x_t^1 = 1 - x_t$, we can obtain the price-dividend ratio p_t if we know the state variable and the wealth-consumption ratios $q_{\theta,t}$.

Finally, we now show that q_0, q_1 solve an ODE system. We guess-and-verify that the equilibrium objects $r_t, \pi_t, \mu_{x,t}, \sigma_{x,t}, p_t, q_{\theta,t}$ are solely functions of x_t (i.e., a Markov equilibrium). Observe that

$$
e^{-(\delta+\frac{\rho}{\gamma})t}(B_t^\theta)^{\frac{1}{\gamma}}S_t^{1-\frac{1}{\gamma}}q_{\theta,t} + \int_0^t e^{-(\delta+\frac{\rho}{\gamma})u}(B_u^\theta)^{\frac{1}{\gamma}}S_u^{1-\frac{1}{\gamma}}du
$$

is a local martingale and must have zero drift. Applying Itô's formula, we have

$$
\begin{aligned}
0 = 1 - \frac{1}{\gamma}\Big[\gamma\delta + \rho + (\gamma-1)r(x) - \frac{1}{2}\frac{1-\gamma}{\gamma}(\theta\beta)^2(1+\chi^2) - \frac{1}{2}\frac{\gamma-1}{\gamma}\|\pi(x)\|^2 - \frac{\gamma-1}{\gamma}(\theta\beta)\big(\begin{smallmatrix}1\\ \chi\end{smallmatrix}\big)\cdot\pi(x)\Big]q_\theta(x)\\
+ \Big[\mu_x(x) + \sigma_x(x)\cdot\Big(\frac{1}{\gamma}\theta\beta\big(\begin{smallmatrix}1\\ \chi\end{smallmatrix}\big) - \frac{\gamma-1}{\gamma}\pi(x)\Big)\Big]q_\theta'(x) + \frac{1}{2}\|\sigma_x(x)\|^2 q_\theta''(x), \quad \theta\in\{0,1\}.
\end{aligned} \tag{C.6}
$$

Equation (C.6) is an ODE for q_θ. Because the equilibrium risk-free rate r and the drift μ_x both depend on $c_{t,t}^\theta/C_t$, hence on p and q_θ, these ODEs are nonlinear and they form a coupled system. As the proposition states, we assume a solution to this system of ODEs, in which case the entire equilibrium is indeed Markovian in x. □

Remark C.1 (Log utility special case)**.** *With log utility ($\gamma = 1$), we have an analytical solution. In particular, from the definition of q_θ, and using the change-of-measure B^θ to convert from $\mathbb{E}$ to $\tilde{\mathbb{E}}^\theta$, we have $q_\theta(x) = (\delta+\rho)^{-1}$. By (C.5), we then have $p(x) = (\delta+\rho)^{-1}$. Plugging these back into (C.4), we have $c_{t,t}^\theta/C_t = \kappa/\delta$ for each θ. The expression for the interest rate and market price of risk become*

$$
\begin{aligned}
r(x) &= \rho + \delta - \kappa + \mu_c - \sigma_c^2 - x\beta\sigma_c\\
\pi(x) &= \sigma_c\big(\begin{smallmatrix}1\\ 0\end{smallmatrix}\big) + x\beta\big(\begin{smallmatrix}1\\ \chi\end{smallmatrix}\big)
\end{aligned}
$$

while the state dynamics become

$$
\begin{aligned}
\mu_x(x) &= \kappa\big[(1-x)\eta^1 - x\eta^0\big] - x^2(1-x)\beta^2(1+\chi^2)\\
\sigma_x(x) &= -x(1-x)\beta\big(\begin{smallmatrix}1\\ \chi\end{smallmatrix}\big)
\end{aligned}
$$

Decomposition of the SDF. Given the expression for the SDF in (48) equating it to the rational agents' marginal utility (or see equation (C.1) for $\theta = 0$), we know that the permanent component of the SDF coincides with the permanent component of $C_t^{-\gamma}$. Using Itô's formula, we can write

$$\frac{C_t^{-\gamma}}{C_0^{-\gamma}} = \exp\Big\{ -\frac{1}{2}(\gamma\sigma_c)^2 t - \gamma\sigma_c \big(\begin{smallmatrix}1\\0\end{smallmatrix}\big) \cdot W_t + \Big(\frac{1}{2}\gamma(\gamma+1)\sigma_c^2 - \gamma\mu_c\Big)t \Big\},$$

and so the martingale component of this process is

$$\frac{H_t}{H_0} = \exp\Big\{ -\frac{1}{2}(\gamma\sigma_c)^2 t - \gamma\sigma_c \big(\begin{smallmatrix}1\\0\end{smallmatrix}\big) \cdot W_t \Big\}.$$

The transitory component of the SDF is thus $G_t = S_t / H_t$, or

$$\frac{G_t}{G_0} = \Big(\frac{1-x_t}{1-x_0}\Big)^{-\gamma} \exp\Big\{ -\Big(\rho + \gamma\mu_c - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2\Big)t \Big\}.$$

Next, we use (49) to define the average belief distortion. Noting that $dB^0/B^0 = 0$, we thus have $dB/B = \eta dB^1/B^1$. From here, we may apply Itô's formula to obtain $d\log B_t = \eta dB_t^1/B_t^1 - \frac{1}{2}\eta^2 d[B^1]_t/(B_t^1)^2 = \eta\beta\big(\begin{smallmatrix}1\\\chi\end{smallmatrix}\big) \cdot dW_t - \frac{1}{2}\eta^2\beta^2(1+\chi^2)dt$. Integrating this result, we have

$$\frac{B_t}{B_0} = \exp\Big\{ -\frac{1}{2}\eta^2\beta^2(1+\chi^2)t - \eta\beta\big(\begin{smallmatrix}1\\\chi\end{smallmatrix}\big) \cdot W_t \Big\}.$$

Defining $\tilde{S}_t := S_t/B_t$ and $\tilde{H}_t := H_t/B_t$, as required by the aggregation results given in Appendix D, we then obtain

$$\begin{aligned}
\frac{\tilde{S}_t}{\tilde{S}_0} &= \exp\Big\{ -\int_0^t \Big(r_u + \frac{1}{2}\|\pi_u\|^2 - \frac{1}{2}\eta^2\beta^2(1+\chi^2)\Big)du - \int_0^t \Big(\pi_u - \eta\beta\big(\begin{smallmatrix}1\\\chi\end{smallmatrix}\big)\Big) \cdot dW_u \Big\} \\
\frac{\tilde{H}_t}{\tilde{H}_0} &= \exp\Big\{ -\frac{1}{2}\Big((\gamma\sigma_c)^2 - \eta^2\beta^2(1+\chi^2)\Big)t - \Big(\gamma\sigma_c\big(\begin{smallmatrix}1\\0\end{smallmatrix}\big) - \eta\beta\big(\begin{smallmatrix}1\\\chi\end{smallmatrix}\big)\Big) \cdot W_t \Big\}
\end{aligned}$$

Volatilities of the SDF components. We use the localized entropy definition given in equation (50) to compute

$$\begin{aligned}
\mathcal{L}_t(S) &= \frac{1}{2}\|\pi_t\|^2 \\
\mathcal{L}_t(H) &= \frac{1}{2}(\gamma\sigma_c)^2 \\
\mathcal{L}_t(G) &= \frac{1}{2}\beta^2(1+\chi^2)x_t^2 \\
\mathcal{L}_t(B) &= \frac{1}{2}\eta^2\beta^2(1+\chi^2) \\
\mathcal{L}_t(\tilde{S}) &= \mathcal{L}_t(S) + \mathcal{L}_t(B) - \eta\beta\big(\gamma\sigma_c + \beta(1+\chi^2)x_t\big) \\
\mathcal{L}_t(\tilde{H}) &= \mathcal{L}_t(H) + \mathcal{L}_t(B) - \eta\beta\gamma\sigma_c
\end{aligned}$$

D Heterogeneity and aggregation

In reality, survey data do not contain a single measure of beliefs, but rather an entire cross-section. A common practice is to compute summary statistics like the cross-sectional average belief about an asset return. If we use such an average belief, then whose marginal utility are we bounding via Lemma 1?

Suppose there are $i \in \{1, \ldots, N\}$ individuals making return forecasts. All the analysis from the previous sections goes through, i-by-i. In particular, each of their beliefs can be characterized by a $\mathbb{P}$-martingale B^i_t, and each of them have a marginal utility process $\tilde{S}^i_t$. Assuming perfect risk-sharing for the shocks that are spanned by asset payoffs, we can treat $\tilde{S}^i_t$ as the projection of true individual marginal utility onto the return space. In that case, we have the generalization of formula (6) to

$$S_t = \tilde{S}^i_t B^i_t, \quad \forall i. \tag{D.1}$$

Following the permanent-transitory factorization arguments as above, we then also have the generalization of formula (9),

$$H_t = \tilde{H}^i_t B^i_t, \quad \forall i, \tag{D.2}$$

where $\tilde{H}^i$ is the permanent component of $\tilde{S}^i$, i.e., $\tilde{H}^i$ is a martingale under $\tilde{\mathbb{P}}^i$. All agents share the same transitory component $\tilde{G}^i = G$.

Averaging beliefs amounts to the following aggregation procedure. Consider B defined by

$$\frac{B_{t+1}}{B_t} := \frac{1}{N} \sum_{i=1}^{N} \frac{B^i_{t+1}}{B^i_t}. \tag{D.3}$$

Then, B is a $\mathbb{P}$-martingale as in the initial analysis. Computing the cross-sectional average belief is equivalent to using B:

$$\frac{1}{N} \sum_{i=1}^{N} \tilde{\mathbb{E}}^i_t[X_{t+1}] = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_t[\frac{B^i_{t+1}}{B^i_t} X_{t+1}] = \mathbb{E}_t[\frac{B_{t+1}}{B_t} X_{t+1}] =: \tilde{\mathbb{E}}_t[X_{t+1}].$$

In that case, to satisfy the original relation (9), we must define

$$\tilde{S}_t := \frac{1}{N} \sum_{i=1}^{N} \left(\frac{B^i_t}{B_t} \right) \tilde{S}^i_t. \tag{D.4}$$

and

$$\tilde{H}_t := \frac{1}{N} \sum_{i=1}^{N} \left(\frac{B^i_t}{B_t} \right) \tilde{H}^i_t. \tag{D.5}$$

One can think of $\tilde{S}$ and $\tilde{H}$ as the "aggregated marginal utility" and its permanent component, and equations (D.4)-(D.5) shows that these are really belief-weighted-averages of individual variables. Furthermore, it is straightforward to check that $\tilde{H}$ is a $\tilde{\mathbb{P}}$-martingale, where $\tilde{\mathbb{P}}$ is defined via B (i.e., through the averaging procedure). With this definition, the same bounds on $\tilde{S}$ and $\tilde{H}$ hold as the previous analysis (e.g., Lemma 1 and subsequent results), i.e., we have

$$\tilde{L}_t\left(\underbrace{\frac{1}{N}\sum_{i=1}^{N}\Big(\frac{B^i_{t+1}/B^i_t}{B_{t+1}/B_t}\Big)\frac{\tilde{S}^i_{t+1}}{\tilde{S}^i_t}}_{=\tilde{S}_{t+1}/\tilde{S}_t}\right) \geq \underbrace{\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbb{E}}^i_t}_{:=\tilde{\mathbb{E}}_t}\left[\log R_{t+1} - \log R^f_t\right] \tag{D.6}$$

$$\tilde{L}_t\left(\underbrace{\frac{1}{N}\sum_{i=1}^{N}\Big(\frac{B^i_{t+1}/B^i_t}{B_{t+1}/B_t}\Big)\frac{\tilde{H}^i_{t+1}}{\tilde{H}^i_t}}_{=\tilde{H}_{t+1}/\tilde{H}_t}\right) \geq \underbrace{\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbb{E}}^i_t}_{:=\tilde{\mathbb{E}}_t}\left[\log R_{t+1} - \log R^\infty_{t+1}\right] \tag{D.7}$$

Bounds (D.6)-(D.7) show that we may average the survey expected returns, but we must interpret this average belief as a lower bound on the volatility of *belief-weighted-averages* of individual-level variables like $\tilde{S}^i$ and $\tilde{H}^i$. Interpreting the bounds introduces some nuances in a heterogeneous-agent world.[21]

[21] An alternative to averaging is to simply use the entire cross-section of the survey. In particular, the following individual-specific bounds hold:

$$\tilde{L}^i_t\Big(\frac{\tilde{S}^i_{t+1}}{\tilde{S}^i_t}\Big) \geq \tilde{\mathbb{E}}^i_t\big[\log R_{t+1} - \log R^f_t\big] \tag{D.8}$$

$$\tilde{L}^i_t\Big(\frac{\tilde{H}^i_{t+1}}{\tilde{H}^i_t}\Big) \geq \tilde{\mathbb{E}}^i_t\big[\log R_{t+1} - \log R^\infty_{t+1}\big], \tag{D.9}$$

where $\tilde{L}^i_t$ is defined analogously to $\tilde{L}_t$. One can measure the right-hand-side via an individual-specific survey response. If an analogous version of UEI holds for $\tilde{S}^i_{t+1}/\tilde{S}^i_t$ and $\tilde{H}^i_{t+1}/\tilde{H}^i_t$, then these bounds in conjunction with (13) and (15) can be used to compare the marginal utility and belief distortion volatilities *for a cross-section of investors*.

E Simple growth-optimal portfolios

With the exception of the machine-learning approach from Table 6, the other portfolios studied in this paper are formed solely as combinations of the aggregate stock market with short-term bonds. One can view these "simple" portfolios as proxies for the growth-optimal portfolio constructed only using those assets. This section discusses the underlying assumptions that would justify these "simple" portfolios.

Assuming a portfolio that solely mixes the stock market and risk-free rate, the growth-optimal problem is

$$\max_{\theta_t} \mathbb{E}_t\left[\log(\theta_t R^m_{t+\Delta} + (1-\theta_t)R^f_t)\right] \tag{E.1}$$

for some re-investment time interval Δ. (Note also that R^f_t here denotes the risk-free rate between t and $t+\Delta$.) Using the lognormal approximation (or second-order approximation to the log), problem (E.1) becomes

$$\max_{\theta_t} \mathbb{E}_t\left[\theta_t R^m_{t+\Delta} + (1-\theta_t)R^f_t\right] - \frac{1}{2}\theta_t^2 \mathrm{Var}_t\left[R^m_{t+\Delta}\right], \tag{E.2}$$

which has the standard mean-variance optimal solution

$$\theta^*_t = \frac{\mathbb{E}_t\left[R^m_{t+\Delta} - R^f_t\right]}{\mathrm{Var}_t\left[R^m_{t+\Delta}\right]} \tag{E.3}$$

Implementation of (E.3) requires proxies for the objective conditional expected excess return and objective conditional variance.

The paper explores one of three sets of assumptions, detailed in the next three subsections. The assumptions are, roughly speaking, as follows:

1. Both expected returns and variances depend similarly on implied volatilities. Under this assumption, we recover the baseline with the market as the growth-optimal portfolio.

2. Expected returns are constant, while volatility is predictable. Under this assumption, we recover "volatility-managed portfolios" as growth-optimal.

3. Both expected returns and variances are constant. Under this assumption, we recover a constant-leverage portfolio as growth-optimal.

E.1 Benchmark assumptions: market as growth-optimal

First, one can use the theory from Martin (2017) to construct expected stock returns with option-implied variance (SVIX). In particular, assume (c.f., equation 15 of Martin, 2017)

$$\mathbb{E}_t[R^m_{t+\Delta} - R^f_t] = R^f_t \mathrm{SVIX}^2_t \times \Delta, \tag{E.4}$$

where SVIX^2_t represents the squared SVIX with the same horizon as the period length.

Assume further that risk-neutral and objective conditional variances coincide, which would arise in conditionally log-normal environments, so that (c.f., equation 13 of Martin, 2017)

$$\mathrm{Var}_t[R^m_{t+\Delta}] = (R^f_t)^2 \mathrm{SVIX}^2_t \times \Delta. \tag{E.5}$$

Under (E.4)-(E.5), we have $\theta^*_t = 1/R^f_t \approx 1$, so that the market is approximately growth-optimal as is effectively assumed in Table 3.

E.2 Volatility-managed portfolios

Second, I refrain from assuming excess returns are predictable (as this is a notoriously difficult exercise), but I do allow for predictable volatility. To implement this in a data-driven way, I run a forecasting model for monthly realized variance. Here, the idea is closer to the "volatility-managed portfolios" model of Moreira and Muir (2017), where higher variance forecasts signal an optimal deleveraging from the stock market.

Let $v_t := 252 \times \left(\frac{1}{N_m}\sum_{i=1}^{N_m}(R^m_{t-i/252})^2 - (\frac{1}{N_m}\sum_{i=1}^{N_m} R^m_{t-i/252})^2\right)$ denote the trailing month realized variance (annualized) constructed from daily returns (N_m is the number of trading days in the month; I use $N_m = 21$). I propose an AR(1) for monthly variance:

$$v_{t+1/12} = \nu + \varrho v_t + \epsilon_{t+1/12}. \tag{E.6}$$

Estimating this model on the aggregate stock market realized variance from 1926:08–2024:05 delivers $\hat{\nu} = 0.013$, $\hat{\varrho} = 0.532$, and an R-squared of 0.2827.[22] Using the one-month variance forecasts $\widehat{\text{Var}}_t[R^m_{t+1/12}] = \frac{1}{12}\left(\hat{\nu} + \hat{\varrho} v_t\right)$ in conjunction with an unconditional estimate (i.e., time-series average) for expected excess monthly returns $\widehat{\mathbb{E}}[R^m - R^f]$, I then construct θ^*_t from (E.3), i.e.,

$$\theta^*_t = \frac{\widehat{\mathbb{E}}[R^m - R^f]}{\hat{\nu} + \hat{\varrho} v_t}. \tag{E.7}$$

As two alternatives, I replace $\hat{\nu} + \hat{\varrho} v_t$ in the denominator of (E.7) with the trailing realized variance, v_t, or a trailing measure of implied variance, e.g., VIX^2_t. When piecing together these monthly investment weights, I form an annual return via

$$R^*_{t+1} = \prod_{i=0}^{11}\left(\theta^*_{t+i/12} R^m_{t+(i+1)/12} + (1 - \theta^*_{t+i/12}) R^f_{t+i/12}\right) \tag{E.8}$$

I construct this alternative growth-optimal proxy.

E.3 Constant-leverage portfolio

Third, I assume that neither returns nor variances are predictable. This leads to

$$\theta^*_t = \frac{\widehat{\mathbb{E}}[R^m - R^f]}{\widehat{\text{Var}}[R^m]}, \tag{E.9}$$

which is an optimal portfolio that has no "market timing" (i.e., no time-variation in the weights) but does embed a leverage adjustment. This approach is taken in Table 4.

[22] I have also run several alternative specifications and found trivial gains in forecasting power. I have augmented the regression with more lags of monthly realized variance. I have tried adding last-month squared-VIX as a regressor. And I have run a GARCH(1,1) model on the monthly demeaned returns. At most, the R-squared rises by 0.01, so I stick with the simple AR(1) model (E.6).

F More details on data

F.1 Data samples and sources

	Frequency	Sample Period
A. Stocks		
Excess return	Monthly	1926:12–2020:12
Nagel-Xu	Quarterly	1972:06–1977:06 and 1987:03–2021:03
Livingston	Semi-annually	1952:05–2021:05
CFO	Quarterly	2000:08–2021:05
B. Treasuries		
Par yield, 1-year	Monthly	1953:04–2024:07
Par yield, 10-year	Monthly	1953:04–2024:07
Par yield, 30-year	Monthly	1961:06–2024:07
ZCB yield, 10-year	Monthly	1971:08–2024:07
ZCB yield, 30-year	Monthly	1985:11–2024:07
Return, 10-year	Monthly	1962:06–2024:07
Return, 30-year	Monthly	1962:06–2024:07
Return, index ($>$ 10-year)	Monthly	1954:04–2020:12
BCFF	Monthly	2001:01–2018:06

Table F.1. Samples from key data sources. Note that the Nagel-Xu data from 1972 to 1977 is annual only.

F.2 Yield data

Figure F.1 displays the very tight relationship between the three data sources I use for par yields. The actual yield series I use takes the FRED yields, then fills any missing values with GSW yields, then fills the remaining missing values with the CRSP yields.

Figure F.2 displays par and zero-coupon bond (ZCB) yields. Some results (Table 1) use separately par and ZCB yields as proxies for long-term bond expected returns, because par yields are available for a longer history (and importantly contain more of the period before 1980 when yields were not secularly declining).

F.3 Bond survey interpolation and bootstrapping

To transform the BCFF survey data (which contains one-year forecasts of 6-month, 1-year, 2-year, 5-year, 10-year, and 30-year par yields) into a full zero-coupon yield curve forecast, I (i) interpolate the par forecasts and then (ii) "bootstrap" a zero-coupon forecast curve.

The interpolation step is a combination of linear interpolation (for maturities $\leq$ 5 years) and a fitted Nelson-Siegel model (for maturities $>$ 5 years). The Nelson-Siegel model says that the cc par yields of maturity n should take the form

$$y(n) = \beta_0 + \beta_1 \frac{1-\exp(-n\theta)}{n\theta} + \beta_2 \Big[\frac{1-\exp(-n\theta)}{n\theta} - \exp(-n\theta)\Big]. \tag{F.1}$$

The parameters $(\beta_0, \beta_1, \beta_2, \theta)$ in equation (F.1) are estimated at every date (thus time-varying) to fit the 2-year, 5-year, 10-year, and 30-year par yield forecasts. I estimate this model by computing the fitting three model-implied slopes: between $n = 5$ and $n = 2$; between $n = 10$ and $n = 5$;

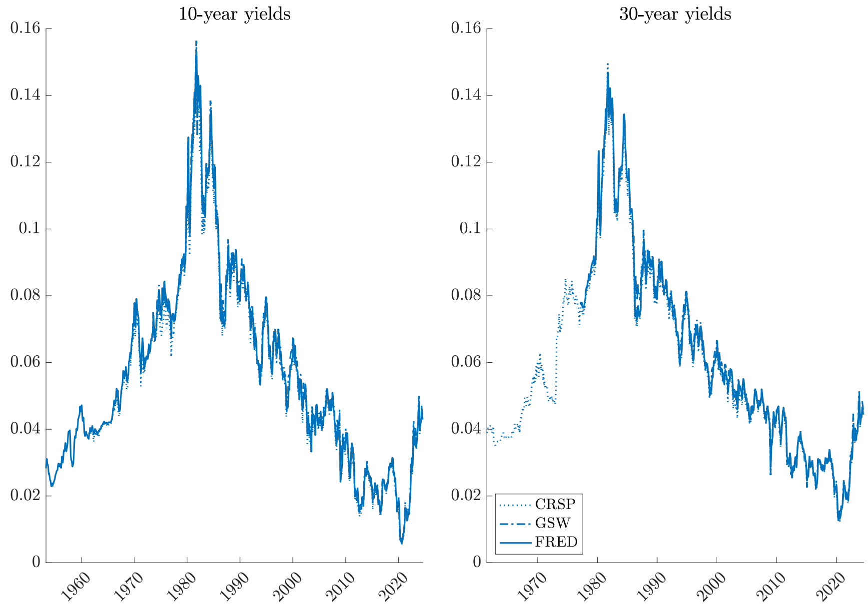

Figure F.1. Actual par yields from the three data sources used in the paper. Yields are annualized and in continuously compounded units (the par yields are converted to cc from coupon-equivalent units).

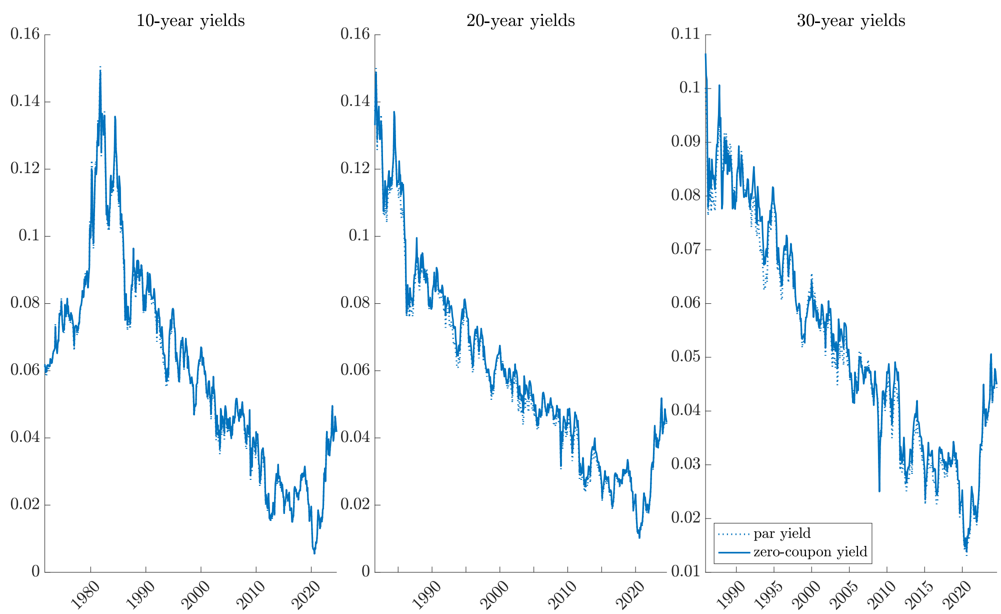

Figure F.2. Actual par and zero-coupon yields from GSW. Yields are annualized and in continuously compounded units (the par yields are converted to cc from coupon-equivalent units).

and between $n = 30$ and $n = 10$. I compute these three slopes both model and survey, after first converting the survey yield forecasts to cc units. Since the slopes difference out β_0, this procedure estimates $(\beta_1, \beta_2, \theta)$ to fit these three slopes. I weight the three slopes by the number of maturity years enveloped in the two maturities (i.e., weights of 3, 5, and 20). I then finally pick β_0 to match exactly the 30-year forecast level, since the long-maturity bonds are the most important in my analysis. Figure F.3 displays the resulting Nelson-Siegel fit to the key par yield forecasts. Figure F.4 contrasts the Nelson-Siegel fit to a linear interpolation in January of each year: the key distinction

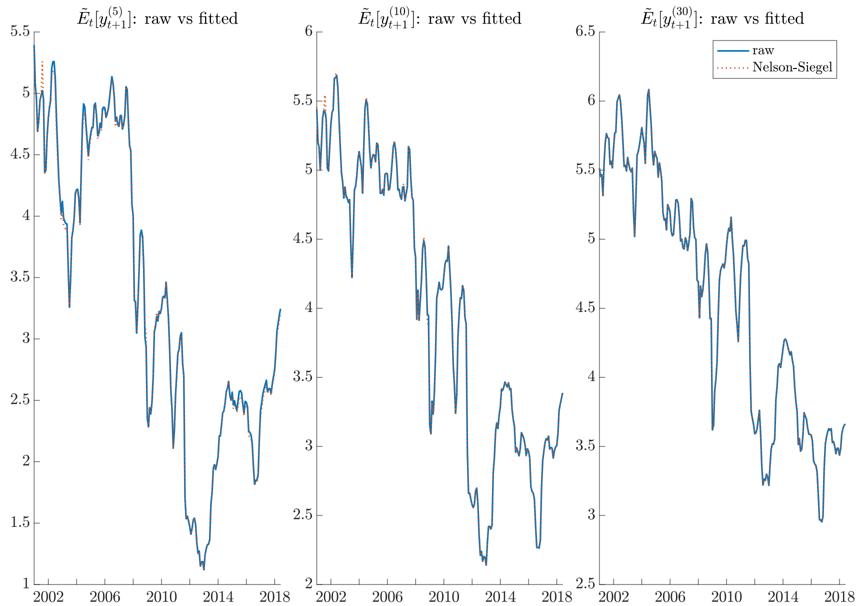

Figure F.3. BCFF consensus expectations of one-year ahead par yields ("raw", solid lines) as well as values from a Nelson-Siegel model fitted to the one-year ahead forecasts (dotted lines). All yield expectations are converted from coupon-equivalent to continuously compounded units.

is the introduction of concavity at the long end, which is typically a feature of the yield curve.

Next, I bootstrap the zero-coupon forecasts as follows. A par yield Y_N^p (in coupon-equivalent units) is the yield such that an N-year bond paying that yield as its semi-annual coupon would trade at par, i.e.,

$$1 = \frac{Y_N^p}{2} \sum_{n=1}^{2N-1} \exp\Big(-\frac{n}{2} y_{n/2}\Big) + \Big(1 + \frac{Y_N^p}{2}\Big) \exp\Big(-N y_N\Big), \tag{F.2}$$

where y_j is the cc zero-coupon yield for maturity j. Thus, one can think of these par yields as the relevant discount rate for a coupon bond. One can solve for the implied N-period zero-coupon yield as

$$y_N = -\frac{1}{N} \log\left[\frac{1 - \frac{Y_N^p}{2} \sum_{n=1}^{2N-1} \exp\big(-\frac{n}{2} y_{n/2}\big)}{1 + \frac{Y_N^p}{2}}\right] \tag{F.3}$$

I recursively obtain $(y_{n/2})_{n=1}^{60}$ in this way. The only nuance is that I treat Y_N^p as the par yield *forecast* and obtain the zero-coupon yield *forecast* y_N, which effectively ignores the nonlinearities in this transformation.

Figure F.5 shows yield curve forecasts (par vs zero-coupon) at the same January dates as above. Notice that the zero-coupon yield forecasts are higher than the par forecasts, mainly because zero-coupon bonds have higher durations due to their absence of coupons. Finally, Figure F.6 displays yield forecast time series and shows how the various transformations affect these forecasts. The left panel shows linearly interpolated par forecasts. The middle panel shows the Nelson-Siegel fitted par forecasts. The right panel shows the bootstrapped zero-coupon yield forecasts.

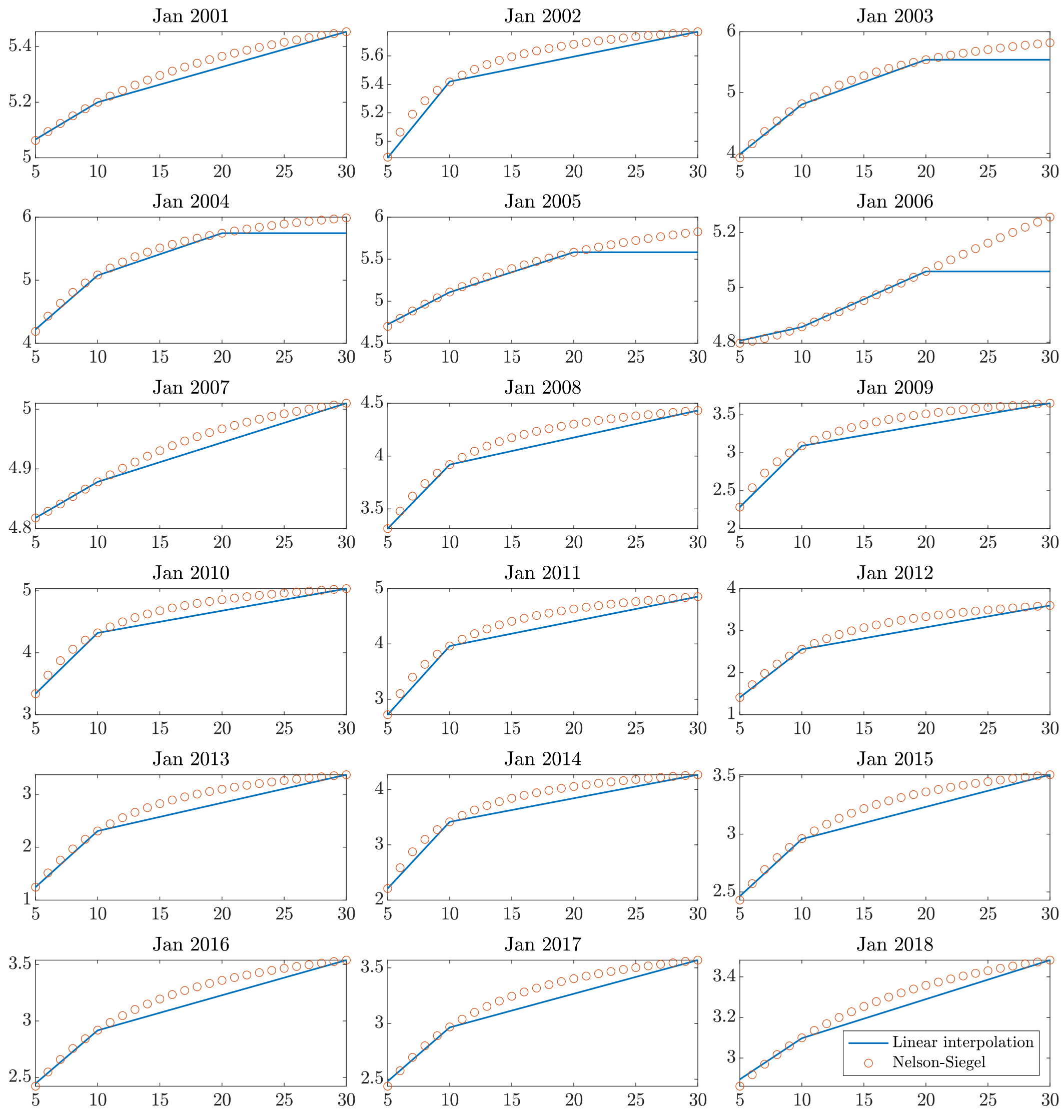

Figure F.4. BCFF consensus expectations of the one-year par yield curve (constructed via linear interpolation of different maturities versus interpolation via fitting a Nelson-Siegel model). All yield expectations are converted from coupon-equivalent to continuously compounded units.

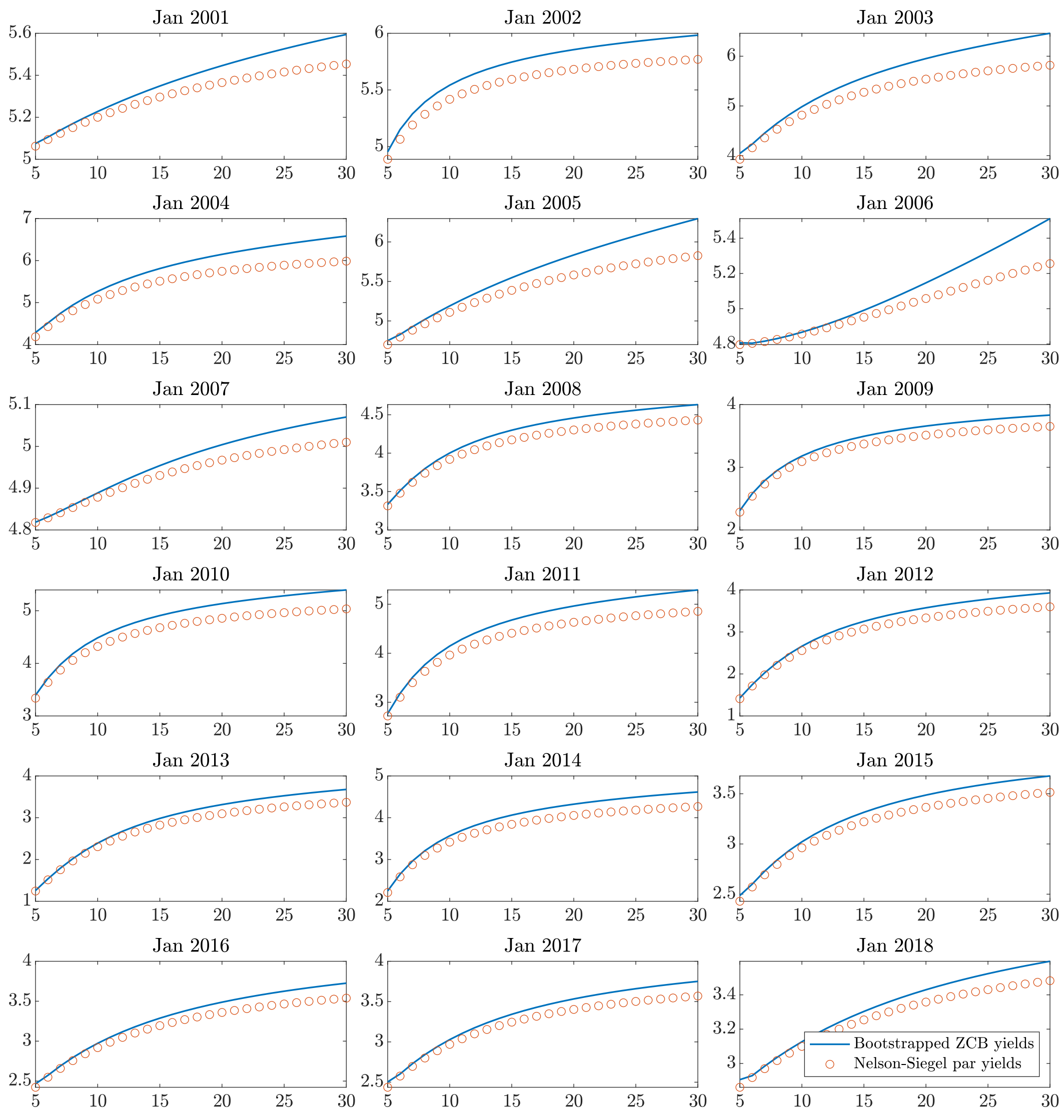

Figure F.5. BCFF consensus expectations of the one-year ahead par yield curve (constructed via fitting a Nelson-Siegel model) and the one-year ahead zero-coupon yield curve (constructed via boostrapping from the Nelson-Siegel par yield curve expectation). All yield expectations are converted from coupon-equivalent to continuously compounded units.

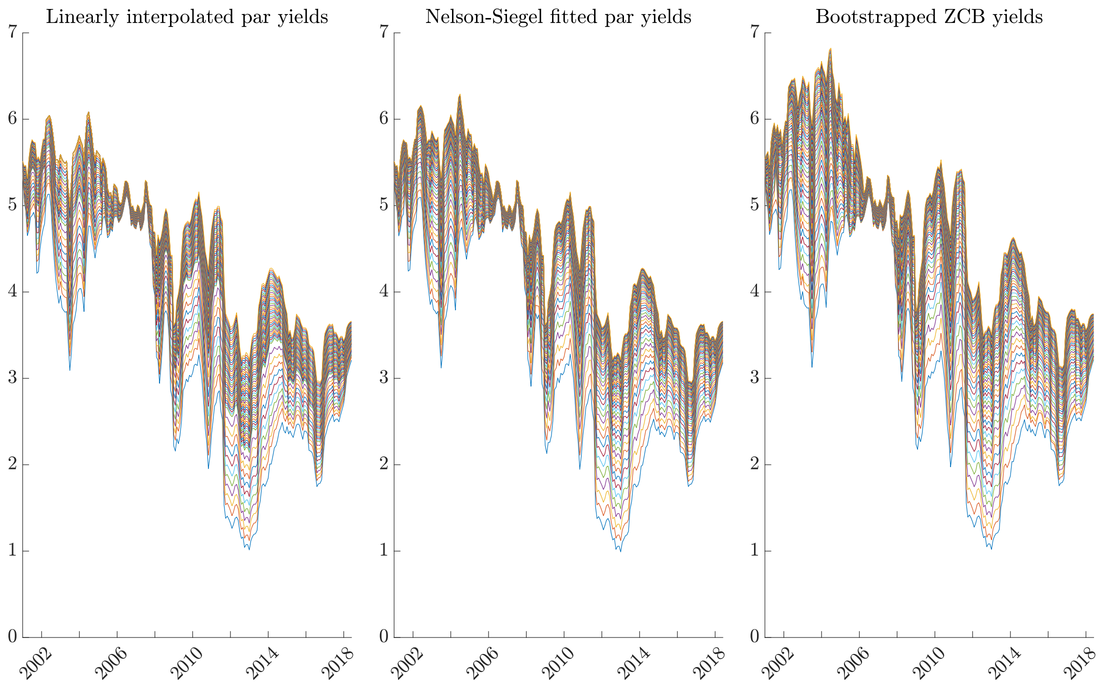

Figure F.6. BCFF consensus expectations of the one-year ahead yield curve. The first panel displays expectations of the par yield curve, constructed by linearly interpolating between maturities. The second panel displays expectations of the par yield curve, constructed by fitting Nelson-Siegel model to the expectations. The third panel displays expectations of the zero-coupon yield curve, constructed by bootstrapping from the Nelson-Siegel par curve. All yield expectations are converted from coupon-equivalent to continuously compounded units.