Belief Distortions and Endogenous Risk Aversion

ADEM ATMAZ
Purdue University

ZIBO ZHOU
Purdue University

June 2026*

Abstract

We develop a dynamic equilibrium model in which belief distortions interact with habit formation to generate heterogeneous and time-varying risk attitudes. Two types of investors have identical preference parameters ex ante but differ in beliefs: biased investors extrapolate from recent aggregate growth, while rational investors hold correct beliefs. Through equilibrium wealth transfers, belief distortions affect the consumption, habits, and risk aversion of rational investors as well. This spillover mechanism generates opposing movements in individual risk aversion and non-monotonic aggregate risk aversion over the business cycle. Consequently, market risk premia and volatility rise not only in contractions but also in strong expansions, patterns for which we provide empirical evidence. The model also reconciles positive cross-sectional and negative time-series relations between wealth and portfolios.

JEL Classifications: G00, G11, G12.
Keywords: Distorted beliefs, habits, time-varying risk aversion, expansions, market risk premium, volatility, portfolios.

*Contacts: aatmaz@purdue.edu and zhou1160@purdue.edu. All errors are our responsibility. Authors have nothing to disclose.

1 Introduction

Beliefs and preferences are central to any economic theory. In asset pricing, they jointly determine investors' portfolio and consumption choices and, through equilibrium, asset prices. Reflecting their importance, a large empirical literature has sought to measure how investors form expectations and how their attitudes toward risk evolve over time. Recent evidence shows that beliefs are often subjective and systematically distorted, with many investors extrapolating from recent outcomes (e.g., Vissing-Jorgensen (2003), Greenwood and Shleifer (2014), Cassella and Gulen (2018), Afrouzi, Kwon, Landier, Ma, and Thesmar (2023)). At the same time, a growing body of evidence finds that risk aversion is time varying and generally countercyclical, consistent with decreasing relative risk aversion (DRRA) preferences (e.g., Cohn, Engelmann, Fehr, and Maréchal (2015), Paravisini, Rappoport, and Ravina (2017), Guiso, Sapienza, and Zingales (2018)). More strikingly, evidence points to substantial heterogeneity in these responses. While most investors become more risk averse during aggregate downturns, a sizable minority becomes more willing to take risk (e.g., Guiso, Sapienza, and Zingales (2018)). Existing theories have difficulty accounting for this heterogeneity because aggregate shocks typically move all investors' risk attitudes in the same direction.

In this paper, we develop a tractable dynamic equilibrium model in which belief distortions interact with habit formation to generate heterogeneous and time-varying risk attitudes, providing a mechanism for understanding these patterns and their implications for asset prices. The economy we consider features two types of investors: rational agents, who hold correct beliefs about aggregate growth, and biased agents, who form extrapolative expectations and become optimistic or pessimistic depending on recent economic conditions. These belief distortions affect investors' portfolio choices and, through equilibrium wealth transfers, alter their consumption paths and habits. Because risk aversion depends on the distance between consumption and habit, these changes generate a central spillover effect: distorted beliefs influence not only the risk aversion of biased investors but also that of rational investors through general equilibrium. To ensure a stationary environment and the persistent coexistence of both investor types, we embed this structure in an overlapping-generations (OLG) economy with "new trees" as in Panageas (2020) and Gârleanu, Panageas, and Zheng (2026).

The model generates rich and state-dependent heterogeneity in risk attitudes over the business cycle. During contractions, rational investors' wealth becomes increasingly positively exposed to aggregate shocks, so their risk aversion rises sharply following further negative shocks. In contrast, biased investors become pessimistic about the economy and reduce their wealth exposure to aggregate risk. As economic conditions deteriorate, this positioning allows them to accumulate wealth and experience declining risk aversion. During expansions, a similar mechanism operates through the opposite belief distortion. Biased investors become more optimistic and increase their wealth exposure to aggregate shocks, while rational investors' wealth exposure can turn negative as the economy expands. The resulting equilibrium wealth transfers again cause the two groups' risk attitudes to move in opposite directions.

At the aggregate level, risk aversion is the consumption-share-weighted harmonic mean of individual risk aversion. Because rational investors typically account for a larger share of aggregate consumption, aggregate risk aversion largely tracks their risk attitudes. As a result, it rises sharply during contractions, as in standard habit models, but it can also rise in strong expansions when rational investors' risk aversion increases. The model thus reconciles average countercyclicality with cross-sectional heterogeneity discussed above: aggregate risk aversion is higher in bad times, yet individual risk attitudes can diverge substantially. To our knowledge, this is the first theoretical framework to account for such heterogeneity in risk attitudes over the business cycle.

These rich risk-attitude dynamics also produce notable asset-pricing predictions. In contrast to standard habit models, both the market risk premium and return volatility are not monotonically countercyclical. Instead, they exhibit a hook-shaped pattern: while both rise sharply during contractions, they can also increase during prolonged expansions. Return volatility is elevated because periods of high aggregate risk aversion are also periods in which stock prices become more sensitive to aggregate shocks. The market risk premium also increases in these periods, although the underlying pricing mechanism differs fundamentally across them. During contractions, negative aggregate shocks are bad pricing shocks in the standard sense: they increase aggregate risk aversion and depress stock prices. Consequently, both the market price of aggregate risk and the covariance between stock returns and aggregate shocks become increasingly positive, generating elevated market risk premia. During pro-

longed expansions, however, positive aggregate shocks themselves become bad pricing shocks because they increase aggregate risk aversion and tend to depress stock prices. As a result, both the market price of aggregate risk and the covariance between stock returns and aggregate shocks become negative. Nevertheless, their product remains positive and economically large, generating elevated risk premia during strong expansions. This key feature is absent from standard habit models and other leading asset-pricing frameworks, which typically imply lower risk premia as economic conditions improve. It is, however, consistent with empirical evidence that "good news" during expansions is "bad news" for stocks (e.g., McQueen and Roley (1993), Boyd, Hu, and Jagannathan (2005), Andersen et al. (2007)).

To assess these predictions, we examine the behavior of the market risk premium and return volatility in the data. Using U.S. aggregate consumption and GDP data, we construct empirical measures of the business-cycle state and relate them to conditional estimates of the market risk premium and return volatility for the S&P 500 index. Consistent with the existing literature, a linear specification suggests that both the market risk premium and return volatility are countercyclical. However, this specification masks an important nonlinearity. In piecewise regressions that allow expansions to have a different slope, both the market risk premium and return volatility stop declining and begin to rise, consistent with the model's prediction that risk premia and volatility can increase in expansions.

The model also delivers novel implications for investors' portfolio choices. In standard partial-equilibrium models with DRRA preferences, wealthier investors are less risk averse and therefore allocate a larger fraction of their wealth to the risky asset. Our model preserves this cross-sectional implication: rational investors, who have correct beliefs, tend to accumulate more wealth, have lower risk aversion, and allocate a larger fraction of their wealth to the risky asset than biased investors, consistent with the empirical evidence in Wachter and Yogo (2010) and Calvet and Sodini (2014). Along the time series, however, the logic can reverse. Because of belief distortions, wealth changes are also associated with shifts in perceived risk compensation, which can dominate the direct DRRA effect of wealth on risk taking. Investors may therefore allocate a smaller fraction of wealth to the risky asset when their wealth rises. This is the case for rational investors throughout the business cycle. They lose wealth but increase their risky portfolio fraction in contractions, gain wealth but reduce their risky fraction

in moderate expansions, and again lose wealth but increase their risky fraction in strong expansions as the risk premium rises once more. A similar negative relation holds for biased investors in most states. The model therefore generates what is often viewed as a puzzling negative time-series relation between wealth and risky portfolio allocation despite DRRA preferences, consistent with empirical evidence in Brunnermeier and Nagel (2008) and Meeuwis (2020).

Overall, this paper makes four main contributions. First, it provides a tractable general equilibrium framework in which belief distortions and habit formation jointly determine heterogeneous and time-varying risk attitudes. Second, it identifies a spillover channel through which biased investors' beliefs affect rational investors' risk aversion, generating opposing responses over the business cycle. Third, it shows that this channel produces high market risk premium and volatility not only in contractions but also in strong expansions, and provides suggestive empirical evidence for this nonlinearity. Fourth, it shows that, in general equilibrium with distorted beliefs, DRRA preferences can generate a negative time-series relation between financial wealth and risky portfolio shares, even though they imply a positive cross-sectional relation between them.

Our paper first relates to time-varying risk aversion and habit formation. Early contributions include Sundaresan (1989), Abel (1990), Constantinides (1990), Detemple and Zapatero (1991), and Ferson and Constantinides (1991). Campbell and Cochrane (1999) develop the benchmark external-habit model, where a slow-moving habit generates countercyclical risk aversion, risk premia, and volatility. Menzly, Santos, and Veronesi (2004) extend the mechanism to multiple securities and predictable dividend growth, and Santos and Veronesi (2010) study its cross-sectional implications.[1] Our framework builds on Menzly, Santos, and Veronesi (2004) but replaces the representative agent with investors who disagree about fundamentals. The belief heterogeneity changes portfolios, wealth transfers, consumption paths, and ultimately risk aversion. Thus the same aggregate shock can move risk attitudes in opposite directions across investors, generating heterogeneous risk aversion and non-monotonic risk premia and volatility. This distinguishes the paper from other heterogeneous-agent habit models such as Santos and Veronesi (2022), where aggregate shocks move agents' risk attitudes in the same direction.

[1] For additional applications of habit formation in asset pricing and macro-finance, see, among others, Jermann (1998), Boldrin, Christiano, and Fisher (2001), Wachter (2006), Verdelhan (2010), and Heyerdahl-Larsen (2014).

Second, our paper contributes to theoretical work on subjective beliefs in asset prices. Motivated by extensive survey evidence, a growing strand of this literature models belief distortions as extrapolative, whereby investors overweight recent outcomes when forming expectations (e.g., Barberis, Shleifer, and Vishny (1998), Choi and Mertens (2013), Hirshleifer, Li, and Yu (2015), Barberis, Greenwood, Jin, and Shleifer (2015, 2018), Atmaz (2022), Jin and Sui (2022)). These models have been successful in explaining a variety of asset-pricing phenomena, including return predictability, excess volatility, and boom-bust dynamics.[2] Our paper differs by introducing endogenous risk attitudes: belief distortions affect not only portfolio choices but also consumption, habits, and risk aversion. This channel is absent from standard extrapolative-belief models and is essential for our main results.

Finally, our paper relates to dynamic asset-pricing models that jointly study heterogeneous beliefs and preferences, including Gallmeyer and Hollifield (2008), Yan (2008), Bhamra and Uppal (2014), Chabakauri (2015), and Veronesi (2019). In these models, investors have constant relative risk aversion (CRRA) preferences and differ ex ante in beliefs and risk-aversion coefficients. In contrast, investors in our model differ ex ante only in beliefs and share the same preference parameters. Heterogeneity in risk aversion emerges endogenously in equilibrium: belief distortions affect portfolio choices and wealth transfers, which generate different consumption paths and habits across investors. This endogenous ex-post risk-aversion heterogeneity is absent from CRRA frameworks, where belief heterogeneity affects portfolios, wealth dynamics, and prices but not investors' risk aversion itself. As we show, this endogenous risk-aversion response is crucial for the model's ability to generate rising risk premia and volatility during expansions, consistent with the patterns we document in the data.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 characterizes the equilibrium and the dynamics of individual and aggregate risk aversion. Section 4 examines the model's asset-pricing implications and provides empirical evidence. Section 5 studies investors' portfolio choices. Section 6 concludes. Appendix A contains the proofs, Appendix B describes the numerical analysis, and Appendix C provides additional empirical robustness results.

[2]See also Simsek (2021) for a recent review of how belief-driven speculation can generate boom-bust dynamics and how macroprudential policies that restrict speculation in booms can improve financial market stability. In this respect, our result that aggregate risk aversion can rise in strong expansions points to an endogenous stabilizing force for speculative risk taking during booms, a channel absent from models with fixed risk attitudes.

2 Model

2.1 Endowment and Investors' Beliefs

We consider a continuous-time, infinite-horizon, pure-exchange endowment economy in which uncertainty is driven by a standard Brownian motion ω_t defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. Aggregate endowment, denoted by Y_t, follows a geometric Brownian motion

$$\frac{dY_t}{Y_t} = \mu dt + \sigma d\omega_t, \tag{1}$$

where the constants μ and σ are the mean and volatility of aggregate endowment growth, respectively.

At each point in time t, all investors observe the current level of aggregate endowment Y_t. They differ, however, in how they form expectations about its future growth. Specifically, investors are of two types, *rational or biased,* depending on whether their beliefs are consistent with the true data-generating process. Rational investors, denoted by $\mathcal{R}$, hold correct beliefs about both the mean and the volatility of aggregate endowment growth. Hence, from their perspective, the endowment process is given by (1), and their subjective probability measure $\mathbb{P}^{\mathcal{R}}$ coincides with the objective one $\mathbb{P}$.

Biased investors, denoted by $\mathcal{B}$, correctly perceive the volatility parameter σ but misperceive the mean growth rate of endowment. To capture their belief distortion in a tractable way, we assume that they hold extrapolative beliefs: their expected future growth, X_t, is an exponentially weighted average of past realized growth rates

$$X_t = \int_{-\infty}^{t} \beta e^{-\beta(t-u)} \frac{dY_u}{Y_u}. \tag{2}$$

The parameter $\beta > 0$ controls the degree of extrapolation: a higher β places greater weight on recent endowment growth and less on distant past growth rates when forming expectations. From the perspective of biased investors, the endowment process therefore evolves as

$$\frac{dY_t}{Y_t} = X_t dt + \sigma d\omega_{\mathcal{B}t}, \tag{3}$$

where $\omega_{\mathcal{B}t}$ is a Brownian motion under their subjective probability measure $\mathbb{P}^{\mathcal{B}}$ with $d\omega_{\mathcal{B}t} = d\omega_t - (X_t - \mu)\,\sigma^{-1}dt$, implying the likelihood ratio

$$L_t = \frac{d\mathbb{P}^{\mathcal{B}}}{d\mathbb{P}^{\mathcal{R}}}\bigg|_t, \qquad \text{with} \qquad \frac{dL_t}{L_t} = \frac{X_t - \mu}{\sigma} d\omega_t. \tag{4}$$

We note that when $X_t > \mu$, biased investors are currently optimistic, expecting above-average endowment growth; when $X_t < \mu$, they are pessimistic.

Given the definition of X_t, its dynamics under $\mathbb{P}$ follow an Ornstein-Uhlenbeck process

$$dX_t = \beta\,(\mu - X_t)\,dt + \beta\sigma d\omega_t.$$

The process is mean-reverting with long-run mean μ, implying that conditionally biased $\mathcal{B}$-type investors are in fact unconditionally unbiased. The parameter β governs both the speed of mean reversion and the sensitivity of biased expectations to aggregate shocks. In particular, because $\beta > 0$, positive aggregate shocks raise X_t, whereas negative shocks decrease it, making their expectations procyclical, akin to the standard Bayesian expectations. For convenience, we normalize X_t and work with the *belief distortion*

$$Z_t = \frac{X_t - \mu}{\sigma}, \tag{5}$$

which is essentially $Z_t = \int_{-\infty}^{t} \beta e^{-\beta(t-u)} d\omega_u$. We note that a sequence of positive (negative) endowment shocks leads to a persistently high (low) value of Z_t. In this sense, Z_t can also be interpreted as an index of current economic conditions: a high positive value of Z_t reflects expansions, whereas a low negative value reflects contractions. For this reason, we also construct the same state variable Z_t even in our benchmark economy in which all investors are rational.[3] Therefore, Z_t serves both as a business-cycle state variable and as a measure of biased investors' distorted beliefs.

Remark 1 (Discussion on beliefs). We would like to highlight that our main results do not hinge on modeling biased beliefs specifically as extrapolative expectations. In particular, the mechanism

[3]See also Santos and Veronesi (2022), who develop a model with fully rational investors and introduce a recession index that is closely related to our process Z_t.

through which beliefs spill over into risk aversion, as shown in Section 3, would also arise under other formulations of belief distortions. We adopt extrapolative beliefs because they provide a tractable, parsimonious, and empirically grounded specification: survey and experimental evidence suggests that investors tend to overweight recent outcomes when forming expectations (e.g., Vissing-Jorgensen (2003), Greenwood and Shleifer (2014), Cassella and Gulen (2018), Afrouzi, Kwon, Landier, Ma, and Thesmar (2023)). This approach has also become increasingly common in theoretical work on distorted beliefs (e.g., Barberis, Greenwood, Jin, and Shleifer (2015, 2018), Atmaz (2022), Jin and Sui (2022)). Moreover, this formulation preserves unconditional unbiasedness: all investors' expectations on average coincide with the true mean, consistent with Adam, Matveev, and Nagel (2021).

2.2 Investors' Preferences

To generate time-varying risk preferences, we adopt a standard habit-preference specification and extend it to allow for heterogeneous beliefs. Each investor of type $i \in \{\mathcal{R},\mathcal{B}\}$ maximizes expected lifetime utility

$$\mathrm{E}^i\left[\int_0^\infty e^{-\rho t}\ln\left(C_{it}-H_{it}\right)dt\right], \tag{6}$$

where $\rho > 0$ is the subjective time-discount rate, C_{it} denotes consumption, H_{it} is the external habit level, and expectations are taken under the investor's subjective probability measure $\mathbb{P}^i$.[4]

Under these preferences, the local curvature of utility function is

$$\gamma_{it} \equiv -C_{it}\frac{\frac{\partial^2}{\partial C_{it}}\ln\left(C_{it}-H_{it}\right)}{\frac{\partial}{\partial C_{it}}\ln\left(C_{it}-H_{it}\right)} = \frac{C_{it}}{C_{it}-H_{it}} = \frac{1}{1-\frac{H_{it}}{C_{it}}} > 1. \tag{7}$$

As in the habit literature, the local curvature, summarized by the surplus-consumption ratio $1-(H_{it}/C_{it})$, captures the investor's attitude toward risk and generates time-varying risk preferences. In particular, as C_{it} approaches the subsistence level H_{it}, the investor becomes increasingly risk averse with respect to fluctuations in consumption.

[4]In our model, habit is external, implying that type-i investors take H_{it} as given when solving their optimization problem. This external habit assumption has been widely used in the habit-preference literature; see, for example, Abel (1999), Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), and Santos and Veronesi (2010).

In the representative-agent setting of Campbell and Cochrane (1999), the surplus-consumption ratio is assumed to be mean-reverting and perfectly *positively* correlated with innovations in aggregate consumption growth. Given our focus, following Menzly, Santos, and Veronesi (2004), we instead model its inverse, γ_{it}, as a mean-reverting process that is perfectly *negatively* correlated with innovations in consumption growth with ex-ante identical preference parameters across investors as

$$d\gamma_{it} = \kappa\left(\overline{\gamma} - \gamma_{it}\right) dt - \zeta\left(\gamma_{it} - \lambda\right)\left(d\ln C_{it} - \mathrm{E}_t^i\left[d\ln C_{it}\right]\right), \tag{8}$$

where $\overline{\gamma}$ is the long-run mean, $\kappa > 0$ is the speed of mean reversion, $\zeta > 0$ measures the sensitivity of γ_{it} to consumption shocks, and $\lambda \geq 1$ sets a lower bound for γ_{it}. Unlike in representative agent models, the relevant consumption process here is type-specific: C_{it} is endogenous and differs across investor types in equilibrium because beliefs are heterogeneous. For analytical convenience, we define

$$\eta_{it} \equiv \zeta\left(\gamma_{it} - \lambda\right)\gamma_{it}^{-1} > 0,$$

which is positively related to γ_{it} and denote the innovations in type-specific consumption growth as

$$d\ln C_{it} - \mathrm{E}_t^i\left[d\ln C_{it}\right] = \upsilon_{it} d\omega_{it}, \tag{9}$$

where the endogenous quantity υ_{it} captures the sensitivity of type-i investor's consumption growth to aggregate endowment shocks. Thus, $\upsilon_{it} > 0$ implies that type-i investor's consumption growth is procyclical, increasing after positive aggregate shocks and decreasing after negative aggregate shocks. This formulation allows us to express the dynamics of γ_{it} compactly as

$$d\gamma_{it} = \kappa\left(\overline{\gamma} - \gamma_{it}\right) dt - \gamma_{it}\eta_{it}\upsilon_{it} d\omega_{it}. \tag{10}$$

With a slight abuse of terminology, we refer to γ_{it} as the investor's *risk aversion*.[5]

[5] As emphasized by Constantinides (1990) and Campbell and Cochrane (1999), the local curvature of the utility function, is not exactly the Arrow-Pratt coefficient of relative risk aversion (RRA), which measures attitudes toward pure wealth gambles. As Menzly, Santos, and Veronesi (2004) highlight in a setting like ours, the local curvature of the utility function is a positive affine transformation of the RRA, so γ_{it} can indeed be interpreted as a measure of risk aversion.

To ensure that our habit specification is economically sensible, we impose, as in Menzly, Santos, and Veronesi (2004), that positive shocks to consumption increase the implied habit level for all feasible values of γ_{it}. Formally, we require $\mathrm{Cov}_t^i\left(dC_{it}, dH_{it}\right) = \mathrm{Cov}_t\left(dC_{it}, dH_{it}\right) > 0$, so that habit co-moves positively with consumption. This condition implies the parameter restriction

$$\zeta \leq (2\lambda - 1) + 2\sqrt{\lambda\left(\lambda - 1\right)}. \tag{11}$$

Under this specification, as in standard habit models, time variation in risk preferences is driven by the slow-moving nature of habit. In particular, in response to aggregate endowment shocks, habits adjust more sluggishly than individual consumption.

2.3 Investors' Life-Cycle

The setting above is sufficient to characterize a general equilibrium with heterogeneous beliefs and habit preferences. However, absent an additional force, rational investors, whose beliefs align more closely with realized shocks on average, gradually accumulate wealth and eventually dominate the economy. The result is a degenerate, non-stationary equilibrium. To obtain a stationary wealth distribution and preserve a persistent role for biased investors, we therefore introduce a tractable overlapping-generations (OLG) structure without altering the preference or belief specifications above.

Our OLG structure follows the perpetual-youth framework of Blanchard (1985), adapted to asset pricing environments by Gârleanu and Panageas (2015). For tractability, we follow Panageas (2020) and Gârleanu, Panageas, and Zheng (2026). As in these papers, newly born investors are endowed with claims to newly arriving "trees," rather than with labor income as in Gârleanu and Panageas (2015), Ehling, Graniero, and Heyerdahl-Larsen (2017), and Simsek (2021).

Specifically, a mass π of investors is born and an equal mass π dies per unit of time. By the law of large numbers, the mass of investors born at time $s < t$ who are still alive at time t is

$$N_t\left(s\right) = \pi e^{-\pi(t-s)},$$

so total population is $\int_{-\infty}^{t} N_t(s)\,ds = 1$. A constant fraction $\alpha_{\mathcal{R}} > 0$ of new entrants is rational, and the remaining fraction $\alpha_{\mathcal{B}} = 1 - \alpha_{\mathcal{R}}$ is biased. Let $C_{it}(s)$, $H_{it}(s)$, and $W_{it}(s)$ denote, respectively, the time t consumption, habit, and wealth of an investor of type $i \in \{\mathcal{R}, \mathcal{B}\}$ born at time s. Aggregate consumption, habit, and wealth of type i across all living cohorts are then given by $C_{it} = \alpha_i \int_{-\infty}^{t} N_t(s)\, C_{it}(s)\, ds$, $H_{it} = \alpha_i \int_{-\infty}^{t} N_t(s)\, H_{it}(s)\, ds$, and $W_{it} = \alpha_i \int_{-\infty}^{t} N_t(s)\, W_{it}(s)\, ds$.

Each newly born cohort is equally endowed with claims to "trees" that arrive contemporaneously with their birth. The dividend at time t from a tree born at time $s < t$ is given by

$$Y_t(s) = \delta e^{-\delta(t-s)} Y_t,$$

where $\delta > 0$ is the depreciation rate of trees over time. This specification implies that the sum of endowment produced by all existing trees is equal to the aggregate endowment $Y_t = \int_{-\infty}^{t} Y_t(s)\, ds$. Newborn agents exchange their endowed trees for shares of a risky stock, which can be interpreted as a basket of all currently existing trees or the "aggregate stock market". We denote the aggregate stock price by P_t, which satisfies the aggregation condition $P_t = \int_{-\infty}^{t} P_t(s)\, ds$, where $P_t(s)$ denotes the time-t value of the tree born at time s, with $P_s(s) = \delta P_s$, by no-arbitrage.[6] The return on the aggregate stock is denoted by

$$\frac{dP_t + Y_t dt - P_t(t)\, dt}{P_t} = \mu_{Pt} dt + \sigma_{Pt} d\omega_t, \tag{12}$$

where the mean return μ_{Pt} and the return sensitivity to aggregate shocks σ_{Pt} are determined endogenously in equilibrium. As (12) shows, the return on the aggregate stock consists of three terms: capital gains/losses dP_t/P_t, the dividend yield Y_t/P_t , and the cost of purchasing the new trees in the economy $P_t(t)/P_t = \delta$.[7] In addition to the risky stock, available for trading is a risk-free asset, which is in zero net supply, and its instantaneous return r_t, is determined endogenously in equilibrium. The aggregate market risk premium is denoted by $\Pi_{Pt} \equiv \mu_{Pt} - r_t$.

[6] In this specification, all trees are perfect substitutes, so their values are proportional to one another at any point in time. Rather than tracking a continuum of tree-specific claims, it is therefore convenient to pool them into an aggregate market portfolio and represent trade through a single risky stock.

[7] The last term arises because an investor who holds the aggregate stock at time t owns claims only to trees already outstanding at time t, as in Panageas (2020) and Gârleanu, Panageas, and Zheng (2026). Over the next instant, new trees enter the economy. Since P_{t+dt} includes the value of both existing trees and newly arrived trees, the return on the time-t aggregate stock must subtract the cost $P_t(t)\, dt$ of purchasing the newly issued claims.

Because investors have no bequest motive, we follow the standard approach in this literature and assume that they surrender their financial wealth to insurance companies upon death in exchange for a lifetime annuity. Specifically, an investor of type i born at time s with financial wealth $W_{it}(s)$ at time t receives annuity payments equal to $\pi W_{it}(s)$ per unit of time. In exchange, the insurance sector collects the investor's wealth at death and breaks even in expectation. This arrangement helps sustain a stationary wealth distribution across overlapping cohorts.

2.4 Optimization

Formally, an investor of type $i \in \{\mathcal{R}, \mathcal{B}\}$ born at time s chooses a consumption process $C_{it}(s)$ and an admissible dynamic portfolio strategy $\phi_{it}(s)$, the fraction of their wealth invested in the risky stock, to maximize

$$\max_{C_{it}(s),\phi_{it}(s)} \mathrm{E}_s^i \left[\int_s^\infty e^{-(\rho+\pi)(t-s)} \ln\left(C_{it}(s) - H_{it}(s)\right) dt \right], \tag{13}$$

subject to the dynamic budget constraint

$$dW_{it}(s) = W_{it}(s)\left[r_t + \phi_{it}(s)\left(\mu_{Pt} - r_t\right)\right]dt + W_{it}(s)\,\phi_{it}(s)\,\sigma_{Pt}d\omega_t + \left[\pi W_{it}(s) - C_{it}(s)\right]dt, \tag{14}$$

with initial wealth $W_{is}(s) = (\delta/\pi) P_s$.

Relative to a non-OLG economy, the mortality rate π effectively raises the subjective time-discount rate from ρ to $\rho+\pi$, reflecting the survival risk. The mortality rate π also appears in the last term of the dynamic budget constraint (14) because investors receive annuity payments while alive. Moreover, given these preferences, the surplus-consumption ratio of type i investor born at time s is $1-(H_{it}(s)/C_{it}(s))$, which is assumed to be equal to the type-aggregated surplus-consumption ratio of $1-(H_{it}/C_{it})$. In other words, within each investor type i, each cohort's share of total consumption equals its share of total habit: $C_{it}(s)/C_{it} = H_{it}(s)/H_{it}$. Under this specification, the risk aversion γ_{it} in (7) is identical across all cohorts of type i, which preserves analytical tractability and ensures aggregate dynamics depend only on type-i level, not time-s cohort-level heterogeneity.[8]

[8]Indeed, in the proof of Proposition 1 in Appendix A, we demonstrate that in equilibrium the dynamics of $C_{it}(s)$ is cohort-invariant under this specification.

3 Equilibrium

In this section, we characterize the equilibrium and study how distorted beliefs and habit formation affect individual and aggregate risk aversion. We show that belief distortions generate equilibrium wealth transfers across investors, causing risk attitudes to move in opposite directions over the business cycle, in line with the evidence in Guiso, Sapienza, and Zingales (2018). In particular, we find that rational investors' risk aversion rises sharply in contractions and can also rise during strong expansions, while biased investors' risk aversion tends to fall in these states as their extrapolative positions are rewarded. At the aggregate level, risk aversion largely follows rational investors and as a result, it is countercyclical on average but can also increase in strong expansions.

The definition of equilibrium is standard. The economy is said to be in equilibrium if the price process and consumption-portfolio strategies are such that (i) each investor chooses her optimal consumption-portfolio strategies given the asset prices, habits, and beliefs; (ii) goods, stock, and bond markets clear. In our frictionless setting, there is one source of uncertainty and one risky asset (plus a riskless asset) to trade, thus, the financial market is dynamically complete. Market completeness ensures the existence of a unique state price density, denoted by ξ_t, and also allows us to apply standard martingale methods (Karatzas, Lehoczky, and Shreve (1987), Cox and Huang (1989)) to solve each investor's optimization problem and derive equilibrium quantities through the goods market clearing condition $C_{\mathcal{R}t} + C_{\mathcal{B}t} = Y_t$.

We solve for a stationary Markov equilibrium characterized by four state variables: the (exogenous) process Z_t, the (endogenous) risk aversion processes $\gamma_{\mathcal{R}t}$ and $\gamma_{\mathcal{B}t}$, and the (endogenous) consumption share of rational investors, denoted by $y_t = C_{\mathcal{R}t}/Y_t$. The consumption share of biased investors then becomes $1 - y_t$. To highlight the interaction between time-varying risk preferences and biased beliefs, we often compare the equilibrium in our baseline economy to a benchmark economy with only rational investors. This benchmark economy corresponds to setting $\alpha_{\mathcal{R}} = 1$, which implies $y_t = 1$ for all t.

The following proposition characterizes the equilibrium consumption share y_t, its dynamics, and the sensitivity of consumption to aggregate shocks υ_{it} for each investor type $i \in \{\mathcal{R}, \mathcal{B}\}$ in our economy with distorted beliefs and habit formation.

Proposition 1 (**Equilibrium consumption share**). *In the economy with distorted beliefs and habit formation, the equilibrium consumption share is given by*

$$y_t = \frac{1}{1 + L_t \frac{\alpha_{\mathcal{B}}}{\alpha_{\mathcal{R}}} \frac{\gamma_{\mathcal{B}t}}{\gamma_{\mathcal{R}t}} \frac{G_{\mathcal{B}t}}{G_{\mathcal{R}t}}}, \tag{15}$$

where $G_{it} = \pi \alpha_i \int_{-\infty}^{t} \left(C_{is}(s)/C_{is}\right) G_{is} ds$, *for* $i \in \{\mathcal{R}, \mathcal{B}\}$. *The dynamics of equilibrium consumption share is given by* $dy_t = \mu_{yt} dt + \sigma_{yt} d\omega_t$, *where*

$$\mu_{yt} = y_t (1 - y_t) \kappa \Big(\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}} - \frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}} \Big) + \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} \sigma_{yt} + \frac{\sigma_{yt}^2}{y_t} + \pi \Big[\alpha_{\mathcal{R}} \frac{C_{\mathcal{R}t}(t)}{Y_t} (1 - y_t) - \alpha_{\mathcal{B}} \frac{C_{\mathcal{B}t}(t)}{Y_t} y_t \Big], \tag{16}$$

$$\sigma_{yt} = y_t (1 - y_t) \frac{\sigma (1 + \eta_{\mathcal{R}t}) \eta_{\mathcal{B}t} - \sigma (1 + \eta_{\mathcal{B}t}) \eta_{\mathcal{R}t} - Z_t}{1 + (1 - y_t) \eta_{\mathcal{R}t} + y_t \eta_{\mathcal{B}t}}. \tag{17}$$

Here, the equilibrium sensitivity of consumption to aggregate shocks is given by

$$\upsilon_{\mathcal{R}t} = \frac{\sigma (1 + \eta_{\mathcal{B}t}) - (1 - y_t) Z_t}{1 + (1 - y_t) \eta_{\mathcal{R}t} + y_t \eta_{\mathcal{B}t}}, \qquad \textit{and} \qquad \upsilon_{\mathcal{B}t} = \frac{\sigma (1 + \eta_{\mathcal{R}t}) + y_t Z_t}{1 + (1 - y_t) \eta_{\mathcal{R}t} + y_t \eta_{\mathcal{B}t}}, \tag{18}$$

and the fraction of endowment consumed by newly born investors is given by

$$\frac{C_{\mathcal{R}t}(t)}{Y_t} = \frac{\delta}{\pi} \frac{1 + \frac{\kappa}{\pi + \rho} \overline{\Gamma}_t}{1 + \frac{\kappa}{\pi + \rho} \frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}}, \qquad \textit{and} \qquad \frac{C_{\mathcal{B}t}(t)}{Y_t} = \frac{\delta}{\pi} \frac{1 + \frac{\kappa}{\pi + \rho} \overline{\Gamma}_t}{1 + \frac{\kappa}{\pi + \rho} \frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}}. \tag{19}$$

where

$$\frac{1}{\Gamma_t} = y_t \frac{1}{\gamma_{\mathcal{R}t}} + (1 - y_t) \frac{1}{\gamma_{\mathcal{B}t}}. \tag{20}$$

In the benchmark economy with only rational investors, the equilibrium consumption share is given by $\overline{y}_t = 1$, *which implies* $\overline{\mu}_{yt} = \overline{\sigma}_{yt} = 0$, $\overline{\upsilon}_{\mathcal{R}t} = \sigma$, *and* $\overline{\Gamma}_t = \gamma_{\mathcal{R}t}$.

In the rational benchmark economy, in equilibrium, the representative investor consumes the aggregate endowment Y_t. Therefore, her consumption rises after a positive aggregate shock and falls after a negative one. In our baseline economy, both investor types tend to consume more when aggregate endowment is higher. However, due to differences in beliefs and risk aversion, their consumption shares are not constant but instead evolve endogenously over time, as (15) shows. Specifically, the variation

in an investor type's consumption share is driven by three terms: belief heterogeneity, captured by the likelihood ratio L_t driven by Z_t, differences in risk aversion, captured by the ratio $\gamma_{\mathcal{B}t}/\gamma_{\mathcal{R}t}$, and heterogeneity in generational weights, captured by the ratio $G_{\mathcal{B}t}/G_{\mathcal{R}t}$.

Differences in beliefs generate portfolio heterogeneity, which in turn creates scope for wealth transfers across investor types. In equilibrium, investors whose beliefs are more aligned with recently realized shocks accumulate relatively more wealth and consume more. In particular, sequences of sustained positive or negative aggregate shocks are more consistent with the extrapolative expectations of biased investors. As a result, during contractions ($Z_t < 0$) and expansions ($Z_t > 0$), these investors become relatively wealthier, and their share of aggregate consumption increases. In contrast, the consumption share of rational investors, y_t, is highest during normal times ($Z_t \approx 0$), when there are no persistent aggregate shocks in either direction. These cyclical shifts in consumption shares translate into heterogeneous habits level, generating endogenously time varying differences in risk aversion across investor types. Holding all else fixed, the equilibrium consumption allocation also tilts toward the investor type with higher risk aversion, since these investors have higher marginal utility of consumption. This interaction between beliefs and habit formation, and the resulting variation in risk aversion, plays a central role in our analysis, as we demonstrate in the following sections.

The finite-variation generational weight terms G_{it} in (15) capture the equilibrium impact of the entry of i-type investors with new endowments. A higher G_{it} indicates that cohort entry disproportionately supports the consumption of type i investors, thereby increasing their type's share of aggregate consumption. These terms arise from the OLG structure and ensure the stationarity of the consumption share process, as reflected in the final component of the drift term μ_{yt} in (16). The continuous entry of both types of investors with identical endowments prevents the consumption share from being absorbed at the boundaries, thereby ruling out long-run dominance by either type. When the consumption share y_t approaches zero, this last term becomes positive and is driven by $C_{\mathcal{R}t}(t)/Y_t$ in (19), pushing the share upward; when y_t approaches one, it becomes negative and is driven by $C_{\mathcal{B}t}(t)/Y_t$ in (19), pulling the share downward.[9] Turning to the diffusion term σ_{yt} in (17), a key feature is that it declines with

[9]We provide the summary statistics for the ergodic distribution of the consumption share, along with those of other state variables in the economy, in Table B2 in Appendix B.

Z_t. As a result, during expansions (contractions), further positive (negative) aggregate shocks tend to reduce the consumption share of rational investors, consistent with the mechanism discussed above.

Proposition 1 also reports the sensitivity of each investor type's consumption to aggregate shocks υ_{it} in (18). These sensitivities play a central role in driving time variation in investors' risk aversion, as shown in equation (10) in Section 2.2. In the rational benchmark economy, this sensitivity is constant and equal to the fundamental volatility of aggregate endowment. As a result, positive aggregate shocks always reduce risk aversion, generating the familiar decreasing relative risk aversion behavior of standard habit models. In our model, by contrast, the sensitivities of both investor types are time-varying and driven by belief distortions: $\upsilon_{\mathcal{R}t}$ decreases with Z_t, whereas $\upsilon_{\mathcal{B}t}$ increases with Z_t. This has a key implication: even rational investors' risk aversion is affected by biased beliefs in the economy. That is, through general equilibrium effects, belief distortions spill over across investors, adding to the time variation in risk aversion beyond standard habit-based mechanisms. This spillover fundamentally alters the dynamics of risk aversion. Following positive (negative) aggregate shocks, risk aversion need not decline (rise) and can instead exhibit pronounced non-monotonic behavior over the business cycle, as illustrated in Figure 1.[10]

The left panel of Figure 1 shows that biased investors' risk aversion falls as Z_t rises in expansions, whereas rational investors' risk aversion rises during sufficiently strong expansions ($Z_t \gg 0$). The mechanism operates through equilibrium wealth exposures. In expansions, optimistic biased investors expect positive aggregate shocks to continue and choose positions that make their wealth positively exposed to such shocks, generating procyclical consumption, i.e., $\upsilon_{\mathcal{B}t} > 0$. Further positive shocks therefore raise their consumption relative to habit, lowering their risk aversion. Rational investors take the other side of this optimism. As optimistic beliefs distort equilibrium prices away from the rational benchmark, rational investors scale back their wealth exposure to aggregate shocks; in sufficiently strong expansions, $\sigma\left(1+\eta_{\mathcal{B}t}\right)/\left(1-y_t\right) < Z_t$, this exposure becomes negative, leading to $\upsilon_{\mathcal{R}t} < 0$.

[10] In Figure 1 and subsequent figures, we plot conditional-bin-mean fitted curves to illustrate the relation between a variable on the x-axis and an endogenous quantity that depends nonlinearly on multiple state variables. Specifically, we divide the simulated support of the variable on the x-axis into evenly spaced bins and compute the average value of the endogenous quantity within each bin. The plotted curve is fitted through these bin means. This approach displays a "conditional expectation" relation while averaging out variation due to other state variables. We also consider alternative approaches, including parametric piecewise OLS specifications and cubic-spline fits to capture nonlinear relations. These alternatives produce similar patterns. More details on the plotting method are provided in Appendix B.

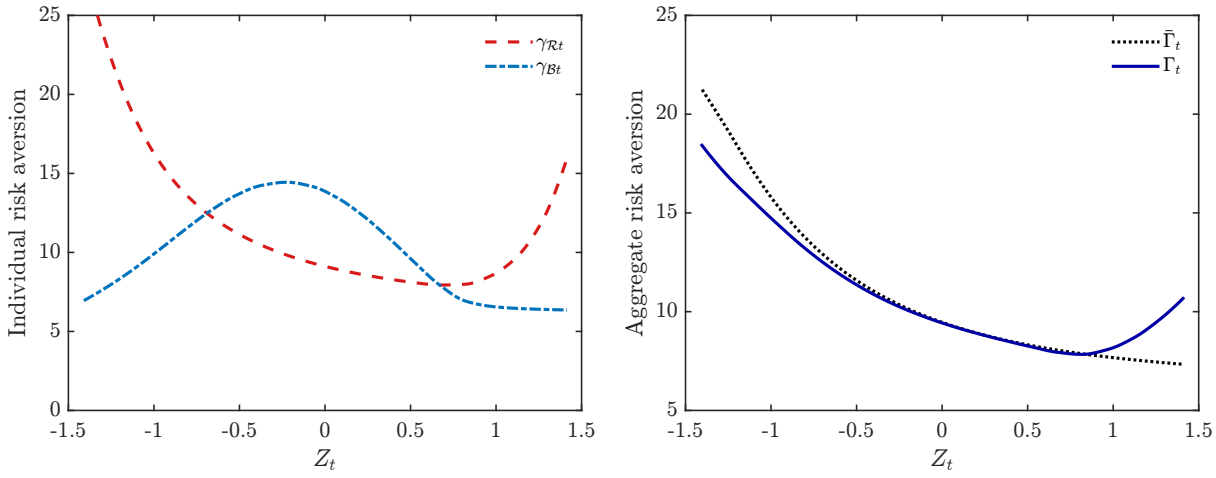


Figure 1. Equilibrium risk aversion. This figure plots equilibrium risk aversion against the state variable Z_t. The left panel plots individual risk aversion of rational investors, $\gamma_{\mathcal{R}t}$, and biased investors, $\gamma_{\mathcal{B}t}$. The right panel plots the aggregate risk aversion in the baseline economy, Γ_t, and in the rational benchmark economy, $\bar{\Gamma}_t$. In both panels, the plotted curves are fitted through conditional bin means. Simulation details and parameter values are reported in Appendix B.

Further positive shocks then reduce their consumption relative to habit, raising their risk aversion. The contrast is even sharper in contractions. As Z_t falls, biased investors become pessimistic and choose positions that make their wealth negatively exposed to aggregate shocks. Further negative shocks then validate their pessimism, raise their consumption relative to habit, and lower their risk aversion. Rational investors again take the other side: they increase their wealth exposure to aggregate shocks, so further negative shocks reduce their consumption relative to habit and sharply raise their risk aversion.

Proposition 1 also reports the quantity Γ_t in (20), which admits a natural interpretation as aggregate risk aversion, as it corresponds to the consumption-share-weighted harmonic mean of individual risk aversion.[11] The right panel of Figure 1 compares aggregate risk aversion in our economy with the representative investor's risk aversion in the rational benchmark. In the benchmark economy, risk aversion declines monotonically with Z_t because positive aggregate shocks raise the representative investor's consumption and thus lower her risk aversion. Comparing this pattern with the risk aversion of rational investors in our economy highlights the spillover effect of distorted beliefs. In addition to generating the sharp rise in rational investors' risk aversion during strong expansions, belief distortions also amplify

[11] Equivalently, $1/\Gamma_t$ can be interpreted as aggregate risk tolerance, given by the consumption-share-weighted average of individual risk tolerances. This form of aggregation is standard in the heterogeneous risk aversion literature. See, for example, Chabakauri (2013) and Bhamra and Uppal (2014).

its increase in downturns relative to the rational benchmark. Intuitively, to counter biased investors' pessimism, rational investors raise their wealth exposure to aggregate shocks well above one in these states, pushing their consumption closer to habit under further negative shocks.

In the baseline economy, however, aggregate risk aversion Γ_t follows a notably different pattern: it declines over most of the state space but rises again during sufficiently strong expansions. This pattern arises in equilibrium because rational investors, whose beliefs align more closely with realized shocks on average, accumulate more wealth and account for a larger share of aggregate consumption on average. Although their consumption share falls in both contractions and expansions, rational investors typically remain the dominant type, with a larger consumption share than biased investors. Aggregate risk aversion therefore largely tracks rational investors' risk aversion, rising during both contractions and sufficiently strong expansions.[12] As we show in Section 4, by summarizing aggregate risk attitudes in the economy, Γ_t plays a central role in determining asset prices and their dynamics.

We emphasize that our findings in this section are broadly consistent with existing empirical evidence on risk attitudes and help reconcile several otherwise puzzling patterns in the data. A growing body of evidence suggests that investors' risk aversion is countercyclical on average: investors are less willing to bear financial risk in bad times than in good times. For example, Bricker et al. (2014) show that households' willingness to take financial risk declined on average during the 2008–2009 recession. Cohn, Engelmann, Fehr, and Maréchal (2015) provide experimental evidence that financial professionals are, on average, less willing to take financial risk in bust states than in boom states. Guiso, Sapienza, and Zingales (2018) find that Italian households' risk aversion increased on average during the 2008 financial crisis. This average pattern also arises in our framework. As shown in Figure 1, Γ_t, which summarizes the average risk attitudes across heterogeneous investors, is higher in contractions than in expansions.

At the same time, the data also show substantial heterogeneity in risk-attitude changes during economic downturns, which is difficult to capture in existing habit models. Guiso, Sapienza, and Zingales (2018) show that, after the 2008 financial crisis, the risk aversion of most, but not all, clients of an Italian

[12] We note that the behavior of aggregate risk aversion, as well as individual risk aversion, is driven by the interaction between belief distortions and habit preferences. The OLG structure is not necessary for these dynamics; it only guarantees a stationary equilibrium by preventing either investor type from eventually dominating the economy. Additionally, rational investors' consumption-share dominance is also robust, even for small values of their population share $\alpha_{\mathcal{R}}$.

bank increased. In particular, they report that 55% of clients experienced an increase in their quantitative measure of risk aversion, 18% remained unchanged, and 27% became less risk averse (see their Table 2). Importantly, they also find that individuals who experienced large financial losses exhibited a substantial increase in risk aversion. Bricker et al. (2014) provide related evidence. They document substantial heterogeneity in household wealth changes during the crisis: although most households lost wealth, a nontrivial fraction experienced wealth gains. Moreover, households with the largest wealth gains became more willing to take financial risk (see their Table 10.9). Our model speaks directly to this heterogeneity. As shown in Figure 1, under endogenous wealth redistribution, investors who gain wealth during contractions become less risk averse, while investors who suffer large losses become substantially more risk averse. To the best of our knowledge, this is the first theoretical framework to explain both the average countercyclical risk aversion and the heterogeneity in risk attitudes across investors.

4 Asset Prices

In this section, we study the implications of distorted beliefs and habit formation for aggregate asset prices. We show that, unlike in standard habit models, both the market risk premium and return volatility exhibit non-monotonic behavior over the business cycle. In particular, while both quantities rise sharply during contractions, they can also increase during sufficiently strong expansions due to the endogenous response of aggregate risk aversion. We further provide empirical evidence consistent with these patterns. We begin by characterizing the equilibrium aggregate stock price.

Proposition 2 (**Equilibrium aggregate stock price**). *In the economy with distorted beliefs and habit formation, the equilibrium aggregate stock price is given by*

$$P_t = \frac{1}{\rho + \pi + \kappa} \left(1 + \frac{\kappa}{\rho + \pi} \frac{\overline{\gamma}}{\Gamma_t} \right) Y_t, \tag{21}$$

where Γ_t *is as in Proposition 1.*

In the benchmark economy with only rational investors, the equilibrium aggregate stock price is given by $\overline{P}_t = \frac{1}{\rho+\pi+\kappa} \left(1 + \frac{\kappa}{\rho+\pi} \frac{\overline{\gamma}}{\overline{\Gamma}_t} \right) Y_t$.

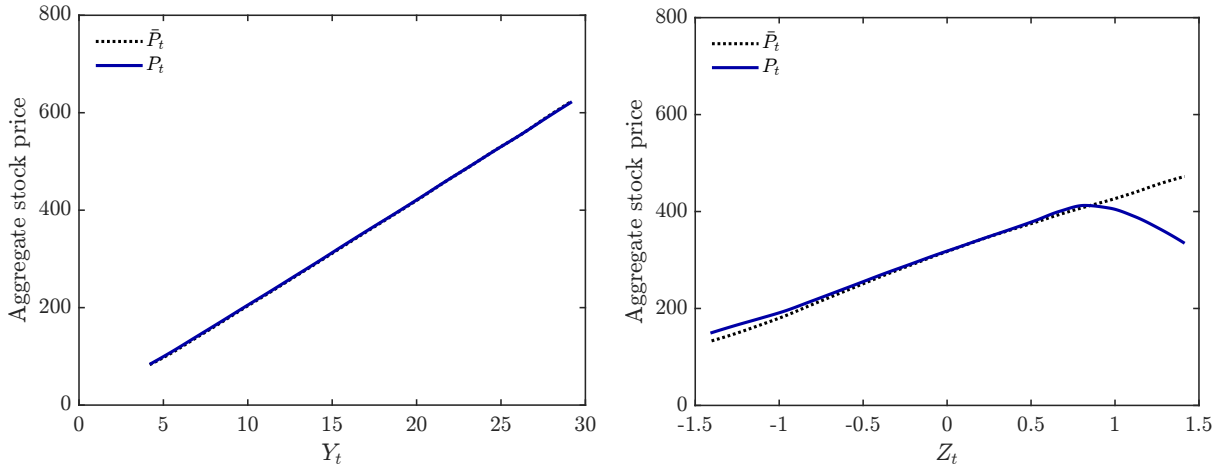


Figure 2. Equilibrium aggregate stock price. This figure plots the equilibrium aggregate stock price in the baseline economy, P_t, and in the rational benchmark economy, $\overline{P}_t$, against aggregate endowment Y_t in the left panel, and against the state variable Z_t in the right panel. In both panels, the plotted curves are fitted through conditional bin means. Simulation details and parameter values are reported in Appendix B.

Proposition 2 shows that, in both benchmark and baseline economies, the aggregate stock price increases with the endowment level Y_t as illustrated in the left panel of Figure 2, and decreases with aggregate risk aversion. The right panel highlights the key difference between the two economies over the business cycle. In the benchmark economy, positive aggregate shocks always reduce rational investors' risk aversion. As investors become more willing to hold the risky stock at the prevailing price, to clear the market, the stock price rises unambiguously during expansions. In contrast, with biased investors present, aggregate risk aversion can increase during strong expansions, as discussed in Section 3. When aggregate risk aversion is higher, investors are collectively less willing to hold the risky stock at the prevailing price, creating downward demand pressure on the aggregate stock price. This downward demand pressure therefore offsets the direct effect of higher endowment and can even reduce the aggregate stock price on average, generating the non-monotonic price pattern. As a result, market return dynamics in our economy can differ from those in the benchmark, particularly during strong expansions.

To characterize these implications, the following proposition derives the equilibrium risk premium and return volatility of the aggregate market.

Proposition 3 (**Equilibrium market risk premium and return volatility**). *In the economy with distorted beliefs and habit formation, the equilibrium return volatility of the aggregate market is given by the absolute value of*

$$\sigma_{Pt} = \sigma + \frac{\frac{\kappa}{\pi+\rho}\frac{\overline{\gamma}}{\overline{\Gamma}}t}{1+\frac{\kappa}{\pi+\rho}\frac{\overline{\gamma}}{\overline{\Gamma}}t}\left[y_t\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t}\frac{\Gamma_t}{\gamma_{\mathcal{R}t}} + (1-y_t)\,\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t}\frac{\Gamma_t}{\gamma_{\mathcal{B}t}} + \sigma_{yt}\left(\frac{\Gamma_t}{\gamma_{\mathcal{R}t}} - \frac{\Gamma_t}{\gamma_{\mathcal{B}t}}\right)\right], \tag{22}$$

and the equilibrium market risk premium by

$$\Pi_{Pt} = \left[\sigma + y_t\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t} + (1-y_t)\,\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t} - (1-y_t)\,Z_t\right]\sigma_{Pt}, \tag{23}$$

where σ_{yt}, Γ_t, and υ_{it} for $i \in \{\mathcal{R}, \mathcal{B}\}$ are as in Proposition 1.

In the benchmark economy with only rational investors, the equilibrium return volatility of the aggregate market is given by $\overline{\sigma}_{Pt} = \sigma + \left[\left(\frac{\kappa}{\pi+\rho}\frac{\overline{\gamma}}{\overline{\Gamma}_t}\right) / \left(1+\frac{\kappa}{\pi+\rho}\frac{\overline{\gamma}}{\overline{\Gamma}_t}\right)\right]\overline{\eta}_{\mathcal{R}t}\overline{\upsilon}_{\mathcal{R}t}$ and the equilibrium market risk premium by $\overline{\Pi}_{Pt} = \left[\sigma + \overline{\eta}_{\mathcal{R}t}\overline{\upsilon}_{\mathcal{R}t}\right]\overline{\sigma}_{Pt}$, with $\overline{\upsilon}_{\mathcal{R}t} = \sigma$.

To highlight how the presence of investors with distorted beliefs alters the behavior of the market return dynamics, we begin with a benchmark economy populated exclusively by rational investors. In this homogeneous setting, rational investors' risk aversion is high during contractions and becomes more sensitive to aggregate shocks, as captured by a higher $\overline{\eta}_{\mathcal{R}t}\overline{\upsilon}_{\mathcal{R}t}$. Greater sensitivity of risk aversion amplifies the response of stock prices to aggregate shocks, generating higher return volatility $\overline{\sigma}_{Pt}$ during contractions, as shown in left panel of Figure 3. At the same time, more risk averse investors become less willing to bear aggregate risks during contractions. Therefore, market clearing requires higher compensation for holding the risky stock, leading to a higher market risk premium, as shown in the right panel of Figure 3.

In our economy, aggregate risk aversion Γ_t is non-monotonic in Z_t. As a result, both return volatility and the market risk premium become non-monotonic over the business cycle, as shown in Figure 3. Return volatility rises during both contractions and sufficiently strong expansions because periods of high aggregate risk aversion are also periods in which Γ_t itself becomes more sensitive to aggregate shocks. As a result, stock prices respond more strongly to shocks, generating elevated volatility. The

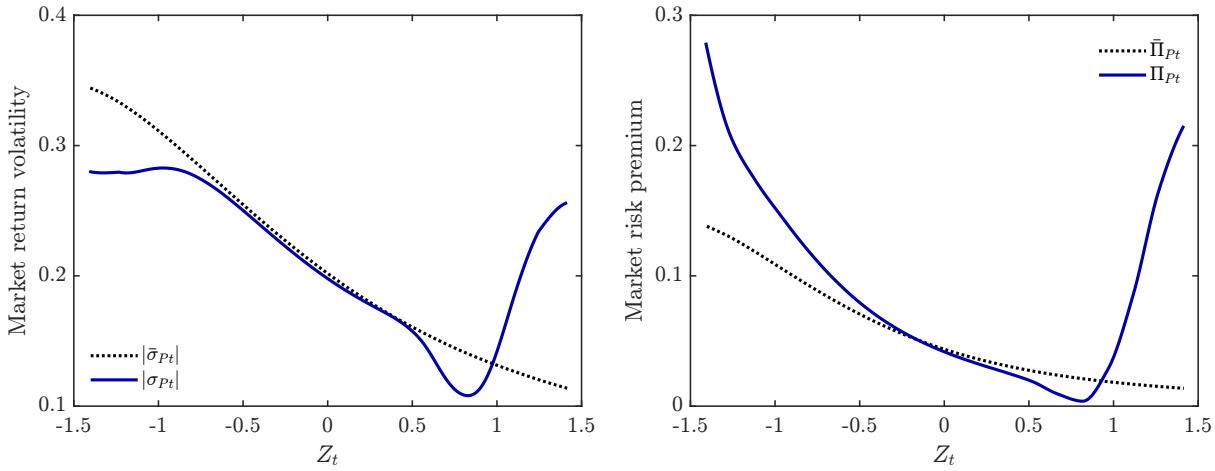


Figure 3. Equilibrium market return volatility and risk premium. This figure plots equilibrium market return volatility in the baseline economy, $|\sigma_{Pt}|$, and in the rational benchmark economy, $|\overline{\sigma}_{Pt}|$, in the left panel and the market risk premium in the baseline economy, Π_{Pt}, and in the rational benchmark economy, $\overline{\Pi}_{Pt}$, in the right panel against the state variable Z_t. The plotted curves are fitted through conditional bin means. Simulation details and parameter values are reported in Appendix B.

market risk premium also increases in both these periods, although the underlying pricing mechanism differs fundamentally across them. During contractions, negative aggregate shocks are bad pricing shocks in the standard sense: they raise aggregate risk aversion and depress stock prices. Consequently, both the return sensitivity σ_{Pt} and the price of aggregate risk (the square bracket term in (23)) become increasingly positive, raising the market risk premium.

During sufficiently strong expansions, however, the mechanism reverses. Positive aggregate shocks now increase Γ_t and tend to depress stock prices, causing positive shocks themselves to become bad pricing shocks. During such times, both the return sensitivity and the price of aggregate risk become increasingly negative as Z_t rises. Nevertheless, their product remains positive and economically large, generating elevated market risk premia on average during these episodes. While a negative market price of risk in good times is a familiar implication of heterogeneous-belief models (see, for example, Panageas (2020)), the negativity of σ_{Pt}, implying that market returns negatively covary with aggregate shocks, $\mathrm{Cov}_t\left[d\ln P_t, d\ln Y_t\right] < 0$, is unique to our framework and arises from elevated aggregate risk aversion during sufficiently strong expansions.[13] This key mechanism is consistent with empirical evidence in

[13]We highlight that the non-monotonic patterns in market risk premia and return volatility shown in Figure 3 are not specific to our baseline calibration. They arise under a broad range of plausible parameter values, including calibrations that use the parameter values of Menzly, Santos, and Veronesi (2004) for the common preference parameters.

McQueen and Roley (1993), who show that when the economy is strong, the stock market responds negatively to news about higher real economic activity because discount rates increase by more than expected cash flows. Similar evidence that "good news" during expansions is "bad news" for stocks is also documented in Boyd, Hu, and Jagannathan (2005) and Andersen, Bollerslev, Diebold, and Vega (2007), among others.

4.1 Empirical Analysis

As shown above, both the market risk premium and return volatility exhibit non-monotonic behavior over the business cycle in our model, in contrast to the conventional view that both decline monotonically as economic conditions improve. In this subsection, we present suggestive empirical evidence supporting these patterns.

We use two commonly employed measures of macroeconomic fundamentals as empirical proxies for aggregate endowment Y_t in the model. The first is aggregate consumption from the Bureau of Economic Analysis (BEA), which is closely aligned with the endowment interpretation in consumption-based asset pricing models. The second is real GDP, also from the BEA, which provides a broader measure of aggregate economic activity. Both measures are at quarterly frequency. For each series, we first construct X_t following the model definition in (2). Specifically, X_t is an exponentially weighted average of past growth rates. For the weights, consistent with both the empirical estimates and our model calibration, we set $\beta = 0.5$ and use the past 30 quarters of consumption or real GDP growth to approximate the infinite-history object X_t.[14] Similarly, following the model definition in (5), we then standardize each resulting X_t series separately using its own sample mean μ and standard deviation σ to construct corresponding $Z_t = (X_t - \mu)/\sigma$.

To evaluate our model implications for market risk premium and return volatility, rather than relying on realized returns, which are noisy proxies for expected returns, we use the annualized conditional risk premium and conditional return volatility series constructed by Schreindorfer and Sichert (2025).[15] The

[14] We also considered alternative values of β that are higher and lower than 0.5 and found that the results remain robust. We further note that increasing the number of lagged quarters beyond 30 has little quantitative effect because the exponential weights decay rapidly over time.

[15] Schreindorfer and Sichert (2025) estimate option-implied risk-neutral return densities from S&P 500 options and recover

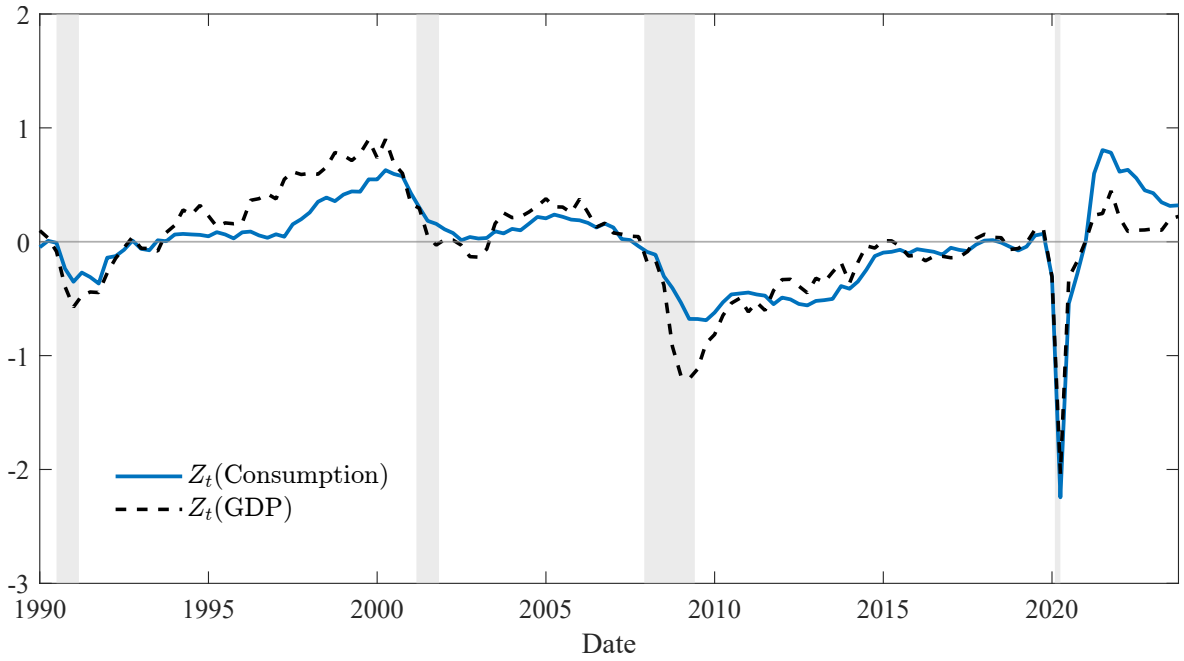


Figure 4. Time series of estimated Z_t. The figure plots the consumption-based Z_t (solid blue line) and real GDP-based Z_t (dashed black line) over the sample period 1990-2023. For each series, we construct Z_t from the corresponding quarterly aggregate growth rates following the model definition. Gray shaded areas denote NBER recession periods.

series are available at a daily frequency, cover the period from 1990 to 2023, and therefore determine our sample period. For robustness, Appendix C considers several alternative data series. First, as proxies for Y_t, we use additional measures to capture business cycle conditions, including industrial production and the real earnings from Robert Shiller's dataset. Second, for return volatility, we use the commonly employed model-free measure based on realized variance from Hao Zhou's website as well as the simpler measure of standard deviation constructed using daily S&P 500 returns. Third, for the market risk premium, we use the option-implied lower bound proposed by Martin (2017), as well as future realized excess returns as a direct but potentially noisy proxy. Finally, because Z_t in the model also reflects belief distortions, we also proxy for it using Michigan survey measures of consumer sentiment and expectations. We show that our results remain robust across these alternative measures.

Figure 4 plots the consumption-based and real GDP-based Z_t series. Both variables summarize

the corresponding physical return densities using a stochastic discount factor specification; the conditional moments of one-month-ahead S&P 500 returns are then computed from these estimated physical densities. We refer to Schreindorfer and Sichert (2025) for details on obtaining the conditional return moments. We thank David Schreindorfer for kindly providing the data.

aggregate economic conditions, with higher values indicating stronger economic activity. The two series are highly correlated, with a correlation of 0.87. Since Z_t is constructed as a decaying weighted average of past growth rates, both series are persistent and move smoothly over time. Given the relatively limited sample size, we define strong expansion states as those in which Z_t lies more than one sample standard deviation above its sample mean, and strong contraction states as those in which Z_t lies more than one sample standard deviation below its sample mean.[16] Both series capture familiar business-cycle episodes: strong contractions occur around the global financial crisis and the Covid pandemic, while strong expansions occur around the dot-com boom. The two measures differ mainly in the post-Covid period: the consumption-based Z_t classifies the rebound as a strong expansion, whereas the GDP-based Z_t does not. This difference reflects the unusually strong consumption-led recovery after the pandemic.

To examine whether the model's implications that the market risk premium and volatility can rise in expansions are supported in the data, we estimate the following continuous piecewise linear regression:

$$\mathcal{Y}_t = \alpha + \beta_1 Z_t + \beta_2 \left(Z_t - c\right)^+ + \varepsilon_t, \tag{24}$$

where Z_t is either consumption-based or real GDP-based state variable, and $\mathcal{Y}_t$ is either the conditional market risk premium or conditional return volatility described above.[17]

Because Z_t measures the business-cycle state in quarter t, we match it with conditional return moments measured at the end of the same quarter. Specifically, because the series are available daily, $\mathcal{Y}_t$ is measured using the conditional moments of one-month-ahead returns on the last trading day of each quarter. This timing choice aligns the information set of $\mathcal{Y}_t$ with the contemporaneous state variable Z_t, as in the model, where both Π_{Pt} and σ_{Pt} are functions of the current state.[18] In the

[16] The sample means of the two Z_t series are close to 0: -0.020 for the consumption-based Z_t and -0.003 for the real GDP-based Z_t, while the corresponding sample standard deviations are 0.380 and 0.437. Under this definition, strong contractions and strong expansions account for about 17% and 13% of the consumption-based Z_t sample, respectively. For the real GDP-based Z_t, the corresponding sample frequencies are about 12% and 13%.

[17] We adopt the continuous piecewise linear regression in equation (24) to match the model-implied continuous relation between asset-pricing quantities and the state variable. Alternative approaches, including simple bin-mean comparisons and independent subsample linear regressions, yield similar qualitative results.

[18] In addition to our baseline last-trading-day construction, we also construct $\mathcal{Y}_t$ using the average conditional moments over the last week of each quarter. To account for potential information frictions, we further consider averages over the week following each quarter-end. The results remain robust across these alternative timing choices.

regression specification (24), the term $(Z_t - c)^+$ equals $Z_t - c$ when $Z_t \geq c$ and zero otherwise. Thus, the slope below the cutoff c is β_1, while the slope above the cutoff is $\beta_1 + \beta_2$. We explore results for two cutoffs. The cutoff $c = 0$ separates broad expansionary states, $Z_t > 0$, from non-expansionary states. The second cutoff is one sample standard deviation above the sample mean of the corresponding Z_t series, $c = \overline{Z} + 1 \times \text{std}$, where $\overline{Z}$ denotes the sample mean of that series. This cutoff isolates strong expansionary states as defined above.[19] Comparing these two cutoffs allows us to ask whether the upward relation is present in economic expansions broadly and whether it becomes stronger in strong expansions, as implied by the model. For comparison, we also estimate the full-sample linear regression

$$\mathcal{Y}_t = \overline{\alpha} + \overline{\beta} Z_t + \epsilon_t. \tag{25}$$

This regression captures the conventional average relation between asset pricing quantities and the business cycle. Comparing (24) and (25) allows us to ask whether a single linear specification masks the high-Z_t reversal predicted by the model. We report the results in Table 1.

Table 1, Panel A, reproduces the well-known univariate linear regression result that both the risk premium and return volatility are countercyclical, i.e., higher in contractions and progressively lower in expansions. However, Panels B and C show that this pattern reverses once we estimate the piecewise regressions in equation (24), which treat expansions separately under the two cutoffs $c = 0$ (Panel B) to $c = \overline{Z} + 1 \times \text{std}$ (Panel C). Across both the consumption-based Z_t and the real GDP-based Z_t, the slope below the cutoff, β_1, is negative for both return volatility and the conditional risk premium, while the high-Z_t slope, $\beta_1 + \beta_2$, is positive. This indicates a positive, rather than negative, relation between asset-pricing quantities and Z_t within expansions. More importantly, comparing $\beta_1 + \beta_2$ across Panels B and C shows that the estimated right-side slopes become substantially larger when the cutoff is raised from $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$. For example, for real GDP-based Z_t, the slopes increase roughly fourfold for return volatility and fivefold for the risk premium, while also becoming statistically more significant. The same pattern also appears for the consumption-based Z_t, especially for the conditional

[19] As discussed in Footnote 16, the sample means of the two Z_t series are close to 0. We therefore use the theoretical cutoff $c = 0$ to distinguish broad expansions from contractions. Using the sample mean $\overline{Z}$ instead of zero in the two cutoff definitions would yield nearly identical results, as we have verified.

	Consumption-based Z_t		GDP-based Z_t	
	Volatility	Risk premium	Volatility	Risk premium
Panel A: full-sample linear regression				
$\overline{\alpha}$	0.142^{***}	0.087^{***}	0.142^{***}	0.087^{***}
	(0.008)	(0.006)	(0.008)	(0.006)
$\overline{\beta}$	-0.012	-0.022	-0.024	-0.028^{*}
	(0.019)	(0.014)	(0.020)	(0.015)
Observations	136	136	136	136
R^2	0.006	0.036	0.033	0.072
Panel B: piecewise regression with cutoff $c = 0$				
α	0.127^{***}	0.076^{***}	0.123^{***}	0.072^{***}
	(0.010)	(0.007)	(0.011)	(0.007)
β_1	-0.053^{**}	-0.055^{***}	-0.073^{***}	-0.066^{***}
	(0.021)	(0.017)	(0.022)	(0.015)
$\beta_1 + \beta_2$	0.057^{*}	0.032	0.050^{**}	0.030
	(0.032)	(0.020)	(0.024)	(0.021)
Obs.$(Z_t \le c)$/Obs.$(Z_t > c)$	61/75	61/75	61/75	61/75
R^2	0.064	0.098	0.125	0.166
Panel C: piecewise regression with cutoff $c = \overline{Z} + 1 \times \text{std}$				
α	0.137^{***}	0.083^{***}	0.134^{***}	0.081^{***}
	(0.008)	(0.006)	(0.007)	(0.005)
β_1	-0.029	-0.040^{**}	-0.054^{***}	-0.051^{***}
	(0.019)	(0.015)	(0.017)	(0.011)
$\beta_1 + \beta_2$	0.128	0.117^{***}	0.204^{***}	0.151^{***}
	(0.081)	(0.040)	(0.047)	(0.050)
Obs.$(Z_t \le c)$/Obs.$(Z_t > c)$	118/18	118/18	120/16	120/16
R^2	0.036	0.086	0.141	0.184

Table 1. Market volatility and risk premium regressions. This table reports regressions of conditional return volatility and the conditional risk premium on the state variable Z_t. The first two columns use the consumption-based Z_t, while the last two columns use the real GDP-based Z_t. Both dependent variables are from Schreindorfer and Sichert (2025) and are measured on the last trading day of each quarter, using the matched quarterly sample from 1990 to 2023. Panel A reports the full-sample linear regression (25). Panels B and C report the continuous piecewise linear regression (24) with cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$, respectively. In the piecewise regressions, the table reports α, the slope β_1 for $Z_t \le c$, and the high-Z_t slope $\beta_1 + \beta_2$ for $Z_t > c$. The row "Obs.$(Z_t \le c)$/Obs.$(Z_t > c)$" reports the number of observations on each side of the cutoff. Newey–West standard errors are reported in parentheses. Statistical significance is denoted by *, **, ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

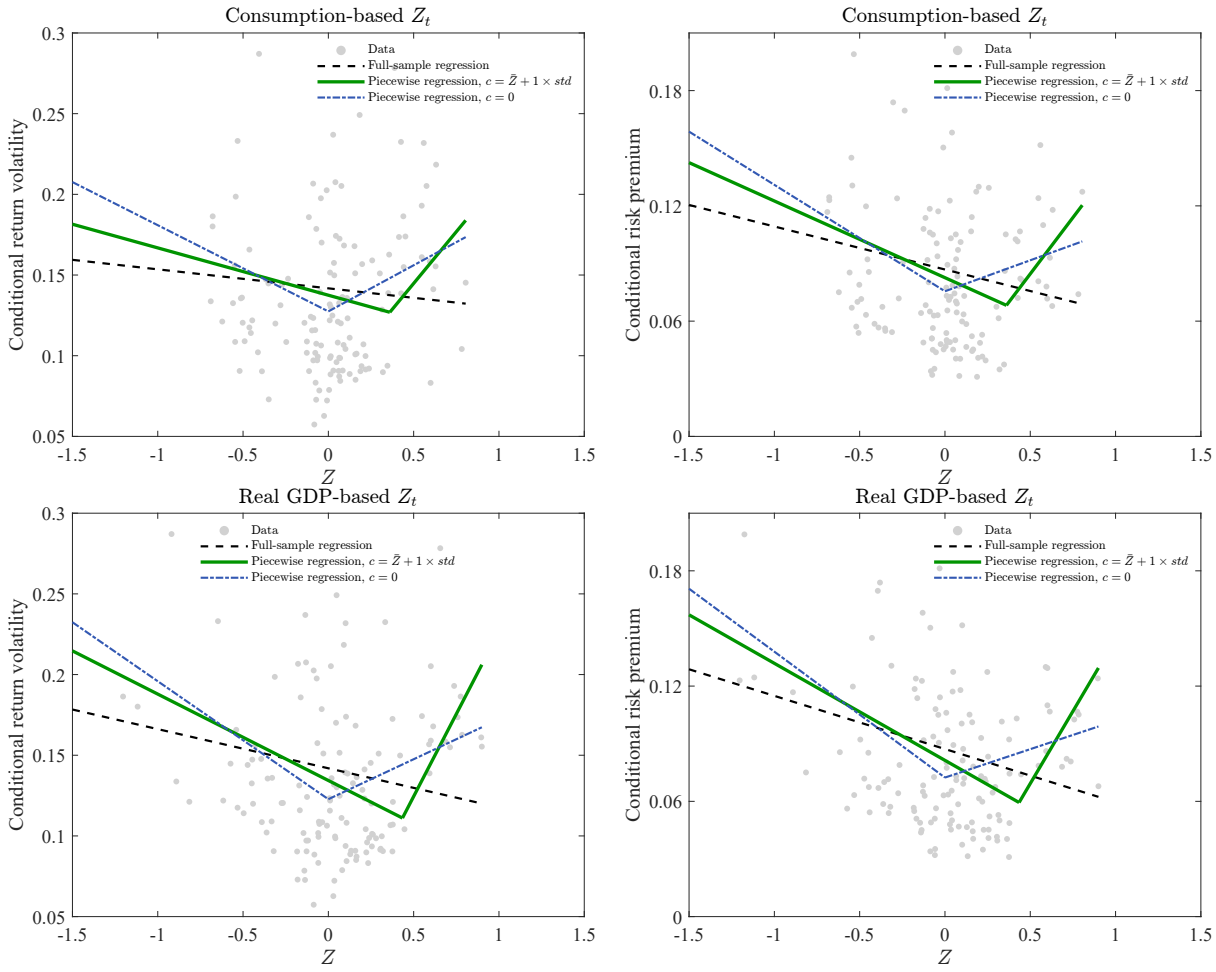


Figure 5. Relation between market return volatility, risk premium, and Z_t. This figure plots conditional market return volatility and risk premium against the estimated Z_t. The top row uses the consumption-based Z_t, and the bottom row uses the real GDP-based Z_t. The left column reports conditional return volatility, and the right column reports the conditional risk premium. Both dependent variables are from Schreindorfer and Sichert (2025) and are measured on the last trading day of each quarter, using the matched quarterly sample from 1990 to 2023. Gray dots denote quarterly observations. The black dashed line is the fitted value from the full-sample linear regression (25). The green solid line is the fitted value from the continuous piecewise linear regression (24) with cutoff $c = \overline{Z} + 1 \times \text{std}$, and the blue dash-dotted line is the fitted value from the continuous piecewise linear regression (24) with cutoff $c = 0$.

risk premium.[20] This finding suggests that the positive relation between asset pricing quantities and Z_t strengthens as Z_t moves further into the right tail, consistent with the model's implication that the spike in volatility and risk premia is concentrated in sufficiently strong expansions.

[20]The statistical inference is based on Newey–West (HAC) standard errors with four quarterly lags, which allow for heteroskedasticity and serial correlation in the regression errors.

We also illustrate these results in Figure 5 by plotting the fitted values from regressions (24) and (25). Consistent with the regression results above, the blue dash-dotted line, based on cutoff $c = 0$, shows that the upward relation is already visible when expansionary states are defined broadly. The green solid line, based on cutoff $c = \overline{Z} + 1 \times \text{std}$, shows that the upward relation becomes much steeper in strong expansions. The black dashed line from the full-sample linear regression instead implies a negative average relation between asset pricing quantities and Z_t, consistent with the conventional view that volatility and risk premia are higher in bad times and lower in good times. Thus, Figure 5 visually reinforces the message from Table 1: the conventional negative relation masks an upturn in expansionary states, especially in strong expansions.

It is worth emphasizing that leading asset pricing models do not generate these patterns for market risk premia and return volatility during strong expansions, even though they provide compelling explanations for why these quantities rise in contractions and adverse states. For example, in habit formation model of Campbell and Cochrane (1999), as well as in our benchmark model, both risk premia and volatility are uniformly countercyclical because expansions are associated with lower marginal utility. In Bansal and Yaron (2004), lower economic uncertainty reduces return volatility and risk premia in good times. In intermediary-based models such as He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014), volatility and risk premia spike when intermediary net worth is low and financial constraints bind, but remain subdued when intermediaries are well capitalized, as typically occur during expansions. In Wachter (2013), volatility and risk premia rise with the time-varying disaster probability, so relatively favorable economic conditions such as expansions imply lower volatility and risk premia. By contrast, our model with distorted beliefs and habit formation can account for elevated market risk premia and volatility in both contractions and sufficiently strong expansions. The key mechanism is endogenous redistribution of wealth, consumption, and risk exposure across investors. This redistribution makes aggregate risk aversion rise not only when the aggregate economy deteriorates, but also during sufficiently strong expansions, generating spikes in both return volatility and market risk premia. To the best of our knowledge, ours is the first work to document this pattern.

5 Investor Portfolios

We now turn from aggregate asset prices to investors' portfolio choices. In partial-equilibrium models with DRRA preferences, investors become more willing to bear risk as their financial wealth rises, implying that the fraction of wealth allocated to the risky asset responds positively to changes in their wealth. Our general-equilibrium setting breaks this direct link. When beliefs are distorted, wealth changes are not only associated with shifts in risk aversion, but also with changes in perceived risk compensation, which can dominate the direct DRRA effect of wealth on risk taking. As a result, even though investors have DRRA preferences, they may allocate a smaller fraction of their wealth to the risky asset following an increase in wealth. This section characterizes this mechanism and shows that the model's portfolio implications are consistent with two empirical patterns: a positive cross-sectional relation between investor wealth and risky portfolio shares (Wachter and Yogo (2010), Calvet and Sodini (2014)), and a negative time-series relation between changes in own financial wealth and risky portfolio shares (Brunnermeier and Nagel (2008), Meeuwis (2020)). We begin by characterizing investors' equilibrium wealth in the following proposition.

Proposition 4 (**Equilibrium wealth**). *In the economy with distorted beliefs and habit formation, the equilibrium wealth of type i investor, for $i \in \{\mathcal{R}, \mathcal{B}\}$, is given by*

$$W_{it} = \frac{1}{\rho + \pi + \kappa} \left(1 + \frac{\kappa}{\rho + \pi} \frac{\overline{\gamma}}{\gamma_{it}}\right) y_{it} Y_t, \tag{26}$$

where $y_{\mathcal{R}t} = y_t$ and $y_{\mathcal{B}t} = 1 - y_t$.

In the benchmark economy with only rational investors, their equilibrium wealth is given by $\overline{W}_{\mathcal{R}t} = \frac{1}{\rho+\pi+\kappa} \left(1 + \frac{\kappa}{\rho+\pi} \frac{\overline{\gamma}}{\overline{\Gamma}_t}\right) Y_t$.

Equation (26) shows that each investor type's wealth is proportional to their consumption, $y_{it} Y_t$. The proportionality factor is the wealth-to-consumption ratio, which decreases with the investor's own risk aversion γ_{it}. Intuitively, when risk aversion is high, current consumption is more valuable because consumption is closer to habit, so investors consume a larger fraction of wealth. To illustrate how investors' wealth evolves over the business cycle, we plot them in Figure 6.

Figure 6 shows substantial wealth redistribution between rational and biased investors over the business cycle. Biased investors' extrapolative beliefs lead them to take positions that benefit from persistent aggregate shocks: positive exposure in expansions and negative exposure in contractions. These positions are rewarded when the economy continues to move in either direction, raising biased investors' relative wealth. Rational investors take the other side of these distorted beliefs, so further negative shocks in contractions and further positive shocks in strong expansions reduce their wealth. Around normal states, where Z_t is close to zero, aggregate shocks exhibit little directional persistence, so biased investors' positions tend to lose money. By contrast, rational investors' beliefs align more closely with realized shocks, allowing them to accumulate substantial wealth. Biased investors' wealth share therefore is the lowest in these states. Although persistent shocks in strong contractions and strong expansions redistribute wealth from rational to biased investors, these episodes are not frequent enough to overturn the wealth accumulated by rational investors in normal states. As a result, rational investors remain wealth dominant throughout the business cycle. These wealth dynamics also illustrate that both investor types exhibit DRRA preferences in equilibrium, since the states in which they hold higher (lower) wealth coincide with their lower (higher) risk aversion, as illustrated in the left panel of Figure 1 of Section 3.

Having determined investors' wealth, we now examine how their portfolios respond to changes in financial wealth. To that end, the following proposition derives their equilibrium portfolios as the fraction of wealth invested in the risky stock.

Proposition 5 (**Equilibrium portfolios**). *In the economy with distorted beliefs and habit formation, the equilibrium portfolios of type i investor, for $i \in \{\mathcal{R}, \mathcal{B}\}$, as a fraction of wealth invested in the risky stock is given by*

$$\phi_{it} = \frac{\Pi_{Pt,i}}{\sigma_{Pt}^2} - \frac{1}{1 + \frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{it}}} \frac{\eta_{it}\upsilon_{it}\sigma_{Pt}}{\sigma_{Pt}^2}, \tag{27}$$

where $\Pi_{Pt,i}$ is the subjective market risk premium of type i investor, with $\Pi_{Pt,\mathcal{R}} = \Pi_{Pt}$ and $\Pi_{Pt,\mathcal{B}} = \Pi_{Pt} + \sigma_{Pt}Z_t$, υ_{it} is as in Proposition 1, Π_{Pt} and σ_{Pt} are as in Proposition 3.

In the benchmark economy with only rational investors, the equilibrium fraction of their wealth invested in the risky stock is given by $\overline{\phi}_{\mathcal{R}t} = 1$.

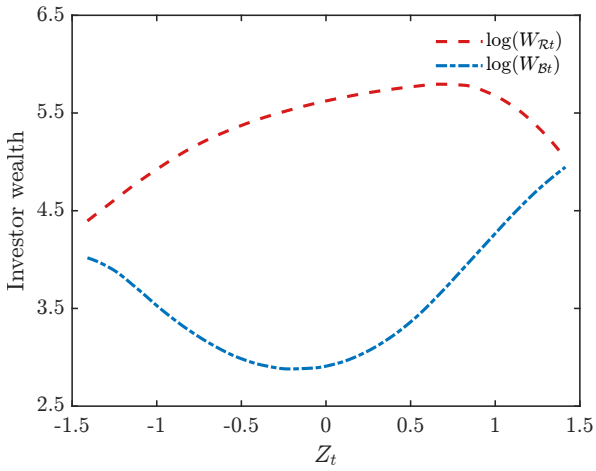


Figure 6. Equilibrium investor wealth. This figure plots equilibrium wealth of rational investors, $\log(W_{\mathcal{R}t})$, and biased investors, $\log(W_{\mathcal{B}t})$, in log scale against the state variable Z_t. The plotted curves are fitted through conditional bin means. Simulation details and parameter values are reported in Appendix B.

In the benchmark economy, market clearing implies that the representative rational investor holds the aggregate market in equilibrium, so the risky portfolio share equals one. By contrast, in our economy, heterogeneity in beliefs and risk aversion leads investors to hold different portfolios and trade in equilibrium. Equation (27) shows that each investor's risky portfolio share (ϕ_{it}) consists of a myopic component and a hedging component. The hedging component, given by the second term in (27), captures investors' demand for protection against fluctuations in future investment opportunities. Because aggregate shocks affect consumption relative to habit, they can shift future risk aversion and hence marginal utility. When an investor's risk aversion positively comoves with the stock returns, as captured by a positive $-\eta_{it}\upsilon_{it}\sigma_{Pt}$, the stock provides a hedge: it pays off when marginal utility is high. This hedging benefit raises the investor's demand for the risky stock.

The myopic component, given by the first term in (27), reflects the usual tradeoff between perceived risk compensation and risk. For rational investors, perceived risk compensation coincides with the objective market risk premium Π_{Pt}, so their myopic demand follows the same non-monotonic pattern of Π_{Pt} discussed in Section 4. For biased investors, by contrast, myopic demand is governed by their distorted subjective risk premium $\Pi_{Pt,\mathcal{B}} = \Pi_{Pt} + \sigma_{Pt}Z_t$, which is plotted in Figure 7. The additional term $\sigma_{Pt}Z_t$ is the wedge between subjective and objective expected stock returns. It captures how biased

expectations about aggregate endowment growth alter expected market returns and can be written as

$$\sigma_{Pt} Z_t = \frac{\mathrm{Cov}_t \left[d \ln P_t, d \ln Y_t \right]}{\mathrm{Var}_t \left[d \ln Y_t \right]} \left(\mathrm{E}_t^{\mathcal{B}} \left[d \ln Y_t \right] - \mathrm{E}_t \left[d \ln Y_t \right] \right) \frac{1}{dt}. \tag{28}$$

Figure 7 shows that biased investors' subjective risk premium rises with Z_t as the economy moves from contractions to moderate expansions, but declines sharply once expansions become sufficiently strong and can even turn negative. This pattern follows from the sign of the stock's exposure to aggregate shocks. During contractions and moderate expansions, stock returns are positively exposed to aggregate endowment shocks, $\mathrm{Cov}_t \left[d \ln P_t, d \ln Y_t \right] > 0$. During these times, extrapolative beliefs have the usual implication: pessimism about Y_t in contractions lowers biased investors' expected stock returns, while optimism in moderate expansions raises them. The mechanism changes in sufficiently strong expansions. As discussed in Section 4, stock returns then become negatively exposed to aggregate shocks, $\mathrm{Cov}_t \left[d \ln P_t, d \ln Y_t \right] < 0$. This creates a notable general-equilibrium effect that differs from standard extrapolative-expectations models (e.g., Barberis et al. (2015), Jin and Sui (2022)): optimism about fundamentals need not imply optimism about asset returns. Although biased investors remain optimistic about the aggregate economy, further positive endowment shocks raise aggregate risk aversion and exert downward pressure on prices. As a result, biased investors expect price depreciation, so their subjective expected stock returns, and hence their subjective risk premium, become negative.

We next examine the cross-sectional relation between investor wealth and risky portfolio shares using simulated data from the model.[21] In the stationary equilibrium of our model, rational investors are wealthier than biased investors as illustrated in Figure 6. This wealth disparity is especially pronounced in normal states, where Z_t is around zero. In these states, favorable realized shocks allow rational investors' consumption to rise well above habit, leading to low risk aversion. By contrast, absent persistent shocks, biased investors lose wealth, pushing their consumption closer to habit and raising

[21] When computing portfolio shares in the simulated sample, we exclude rare observations in which $|\sigma_{Pt}|$ is very close to zero. In these states, the portfolio fraction in (27) can become numerically large because it scales by σ_{Pt}^2. In the baseline analysis, we use the cutoff $|\sigma_{Pt}| < 0.005$, which removes about 0.15% of the stationary sample. This trimming is applied only to the portfolio analysis; all asset-pricing and risk-aversion results above use the simulated sample discussed in Appendix B. The results are robust to alternative cutoffs, including (0.001) and (0.01). We also note that there is no arbitrage in our model as $\sigma_{Pt} \to 0$, since the market risk premium also vanishes when the stock becomes locally riskless.

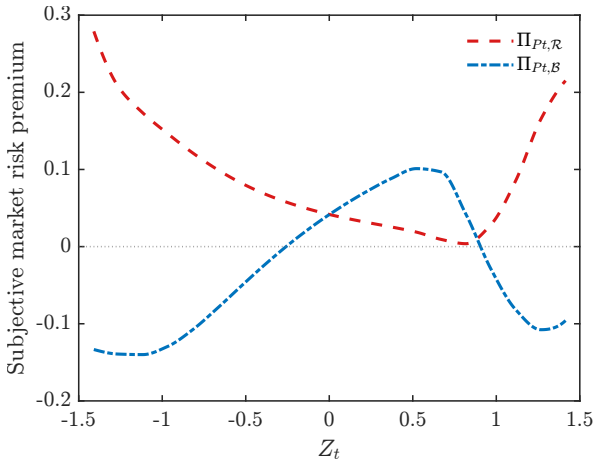


Figure 7. Equilibrium subjective market risk premium. This figure plots equilibrium subjective market risk premia perceived by rational investors, $\Pi_{Pt,\mathcal{R}}$, and biased investors, $\Pi_{Pt,\mathcal{B}}$, against the state variable Z_t. The plotted curves are fitted through conditional bin means. Simulation details and parameter values are reported in Appendix B.

their risk aversion. Consistently, as shown in Table B2 in Appendix B, the median risk aversion of rational investors is 9.18, compared with 12.88 for biased investors. The corresponding sample means are 9.83 and 12.13, respectively. Through standard risk sharing, less risk averse rational investors absorb more aggregate risks and therefore lever up to hold a larger risky portfolio share in normal states. The median portfolio share of rational investors is 1.01, compared with 0.77 for biased investors. This result is consistent with the empirical evidence in Wachter and Yogo (2010), who show that, within each age group, stockholders' median risky portfolio shares increase with net worth. It also broadly aligns with the cross-sectional wealth-portfolio relation documented in Calvet and Sodini (2014), who show that wealthier twins allocate more wealth to risky assets.

We next examine the model's time-series implication for the relation between portfolio adjustments and wealth changes. Motivated by the portfolio-wealth updating relation studied in Wachter and Yogo (2010), we use our baseline monthly simulated observations and estimate

$$\phi_{i,t+1} - \phi_{i,t} = \alpha_i + \beta_i \log\left(\frac{W_{i,t+1}}{W_{i,t}}\right) + \varepsilon_{i,t+1}, \tag{29}$$

where $\phi_{i,t}$ is investor i's risky portfolio share and $W_{i,t}$ is her financial wealth at time t. The coefficient β_i measures how investor i's portfolio share changes with her financial wealth. A negative β_i means

Coefficients	Rational investors	Biased investors	All investors
α	0.004^{***}	0.002^{*}	0.003^{***}
	(0.000)	(0.001)	(0.000)
β	-1.918^{***}	-0.911^{***}	-1.219^{***}
	(0.059)	(0.092)	(0.069)

Table 2. Portfolio adjustment regressions. This table reports pooled regressions of changes in risky portfolio fractions on log changes in financial wealth. The regression specification is in (29). The first two columns report type-specific regressions for rational and biased investors. The last column, "All investors," stacks rational and biased investors in the same sample and estimates a common portfolio-wealth updating coefficient. The regressions use all valid monthly transitions over the 40-year sample window after applying the portfolio-trimming rule. Newey-West standard errors are reported in parentheses and are computed using the optimal lag length $L = 4\,(N/100)^{2/9}$, where N is the number of valid observations in the corresponding regression. Statistical significance is denoted by *, **, ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively. Simulation details and parameter values are reported in Appendix B.

that investors reduce their risky portfolio share when financial wealth increases and increase it when financial wealth decreases. In the type-specific regressions, we use the portfolio and wealth series of rational and biased investors separately. We also stack rational and biased investors in the same sample and estimate a common portfolio-wealth updating coefficient. We refer to this specification as the "All investors" regression. This specification is a natural model analogue of empirical panel regressions that contain many investors in the same sample. We report our findings in Table 2.

As shown in Table 2, the estimated β is significantly negative in all three specifications. Thus, along the time series, investors on average reduce their risky portfolio share when their financial wealth increases and increase it when their financial wealth decreases. The effect is stronger for rational investors than for biased investors and remains strong when both investor types are combined. This negative relation can be understood by comparing investors' financial wealth dynamics with the variation in their perceived risk premia, which drive their portfolio choices over the business cycle.

In contractions, negative shocks shift financial wealth from rational investors to biased investors as Z_t falls. At the same time, rational investors who perceive higher risk premium increase their portfolio shares, while more pessimistic biased investors short sell more, generating a negative relation for both types. In moderate expansions, positive shocks increase financial wealth for both types. Although more optimistic biased investors increase their risky portfolio shares, generating a positive relation, rational investors expect lower future returns and therefore reduce their risky stock exposure, again generating

a negative relation. In strong expansions, the relation turns negative for both types. Rational investors lose wealth but build up their risky portfolio shares as the objective risk premium spikes. Biased investors gain wealth but reduce their stock exposure, and can even take short positions, as their perceived risk premium declines and may become negative. Since the relation is uniformly negative for rational investors and is also negative for biased investors outside moderate expansions, aggregating across states delivers an average negative relation. Therefore, in addition to addressing the cross-sectional portfolio-wealth relation, our model also speaks to the otherwise puzzling time-series evidence in Brunnermeier and Nagel (2008) and Meeuwis (2020), who find a negative response of households' risky portfolio shares to changes in their financial wealth.

Taken together, our findings in this section suggest that the cross-sectional and time-series relations between wealth and portfolio choice need not have the same sign. Cross-sectionally, wealthier investors typically hold larger risky shares, consistent with standard intuition under DRRA preferences. In the time series, however, aggregate shocks can move financial wealth and risky shares in opposite directions due to general-equilibrium effects arising from the interaction between distorted beliefs and risk aversion.

6 Conclusion

This paper develops a dynamic equilibrium model in which belief distortions interact with habit formation to generate heterogeneous and time-varying risk attitudes. Investors have identical preference parameters ex ante, but differ in beliefs: rational investors hold correct beliefs about aggregate growth, while biased investors form extrapolative expectations. These belief distortions affect portfolio choices and equilibrium wealth transfers, which in turn reshape investors' consumption paths and habits. Because risk aversion depends on consumption relative to habit, heterogeneity in beliefs endogenously generates heterogeneity in risk aversion.

The central mechanism is a risk-aversion spillover: distorted beliefs affect not only the risk aversion of biased investors but also that of rational investors through general equilibrium. This mechanism

generates opposing movements in individual risk aversion over the business cycle and makes aggregate risk aversion non-monotonic. As a result, market risk premia and return volatility rise not only in contractions, as in standard habit models, but also in sufficiently strong expansions. In these states, positive aggregate shocks can raise aggregate risk aversion and depress stock prices, so the market price of risk and the covariance between stock returns and aggregate shocks both turn negative, while their product remains positive and large. The model also delivers novel portfolio implications. It preserves the positive cross-sectional relation between wealth and risky portfolio shares implied by DRRA preferences, but can generate a negative time-series relation between wealth changes and risky portfolio adjustments when beliefs are distorted. These predictions help connect the theory to empirical evidence on risk attitudes, asset prices, and household portfolios.

The analysis suggests several directions for future work. On the theoretical side, the framework could be extended to richer belief dynamics, multiple risky assets, production, or intermediary sectors, allowing belief-driven risk-aversion spillovers to interact with cash-flow risk, investment, and financial constraints. On the empirical side, the model highlights the importance of distinguishing expansions and contractions in the study of risk premia and volatility, and the relation between wealth and portfolio choice.

Appendix A: Proofs

Proof of Proposition 1. To determine the equilibrium consumption share and its dynamics in the complete market economy with distorted beliefs and habit formation, we first solve each type of investor's optimization problem using the standard martingale methods. The static problem of type $i = \mathcal{R}, \mathcal{B}$ investor born at time s under the objective measure is given by

$$\max_{C_{it}(s)} \mathrm{E}_s \left[\int_s^\infty e^{-(\rho+\pi)(t-s)} \frac{L_{it}}{L_{is}} \ln \left(C_{it}(s) - H_{it}(s) \right) dt \right], \tag{A.1}$$

subject to

$$\mathrm{E}_s \left[\int_s^\infty e^{-\pi(t-s)} \frac{\xi_t}{\xi_s} C_{it}(s)\, dt \right] = \frac{1}{\pi} \mathrm{E}_s \left[\int_s^\infty \frac{\xi_t}{\xi_s} Y_t(s)\, dt \right] = \frac{\delta}{\pi} P_s, \tag{A.2}$$

where $L_{\mathcal{R}t} = 1$ and $L_{\mathcal{B}t} = L_t$, and ξ_t is the equilibrium state price density. The first-order conditions for the static problem in (A.1)—(A.2) imply

$$C_{it}(s) = \ell_i(s)^{-1} e^{-\rho(t-s)} \left(\frac{\xi_t}{\xi_s} \right)^{-1} \frac{L_{it}}{L_{is}} \gamma_{it}, \tag{A.3}$$

where $\ell_i(s)$ is the Lagrange multiplier associated with the static budget constraint (A.2). From the above expression, we obtain the optimal consumption growth of an i-type investor in cohort s as

$$\frac{C_{it}(s)}{C_{is}(s)} = e^{-\rho(t-s)} \left(\frac{\xi_t}{\xi_s} \right)^{-1} \frac{L_{it}}{L_{is}} \frac{\gamma_{it}}{\gamma_{is}}. \tag{A.4}$$

The aggregate consumption of all investors of type i, $C_{it} = \alpha_i \int_{-\infty}^t N_t(s)\, C_{it}(s)\, ds$, is therefore given by

$$C_{it} = \alpha_i \int_{-\infty}^t N_t(s)\, C_{is}(s)\, e^{-\rho(t-s)} \left(\frac{\xi_t}{\xi_s} \right)^{-1} \frac{L_{it}}{L_{is}} \frac{\gamma_{it}}{\gamma_{is}} ds. \tag{A.5}$$

Next, we define the finite-variation generational weight process G_{it} by

$$G_{it} = \int_{-\infty}^t \pi e^{(\rho+\pi)s} C_{is}(s)\, \xi_s L_{is}^{-1} \gamma_{is}^{-1} ds, \tag{A.6}$$

so that

$$C_{it} = \alpha_i e^{-(\rho+\pi)t} L_{it} \gamma_{it} G_{it} \xi_t^{-1}. \tag{A.7}$$

Hence, using the definition of the rational investors' consumption share $y_t = C_{\mathcal{R}t}/Y_t$, we obtain

$$y_t = \frac{\alpha_{\mathcal{R}} \gamma_{\mathcal{R}t} G_{\mathcal{R}t}}{\alpha_{\mathcal{R}} \gamma_{\mathcal{R}t} G_{\mathcal{R}t} + \alpha_{\mathcal{B}} L_t \gamma_{\mathcal{B}t} G_{\mathcal{B}t}} = \frac{1}{1 + L_t \frac{\alpha_{\mathcal{B}}}{\alpha_{\mathcal{R}}} \frac{\gamma_{\mathcal{B}t}}{\gamma_{\mathcal{R}t}} \frac{G_{\mathcal{B}t}}{G_{\mathcal{R}t}}}, \tag{A.8}$$

as in (17).

To derive the equilibrium dynamics of the consumption share y_t, we next determine dL_t, $d\gamma_{it}$, and dG_{it} for $i = \mathcal{R}, \mathcal{B}$. From (4) and (5), we immediately have $dL_t = L_t Z_t d\omega_t$. The dynamics of endogenous risk aversion γ_{it} are given by (10), where the endogenous consumption sensitivity υ_{it} will be determined below. Applying the Leibniz rule to (A.6) yields

$$\frac{dG_{\mathcal{R}t}}{G_{\mathcal{R}t}} = \alpha_{\mathcal{R}} \pi \frac{C_{\mathcal{R}t}(t)}{Y_t} y_t^{-1} dt, \qquad \text{and} \qquad \frac{dG_{\mathcal{B}t}}{G_{\mathcal{B}t}} = \alpha_{\mathcal{B}} \pi \frac{C_{\mathcal{B}t}(t)}{Y_t} (1 - y_t)^{-1} dt. \tag{A.9}$$

Applying Itô's Lemma to the consumption share in (A.8), using dL_t, $d\gamma_{it}$, dG_{it}, we obtain $dy_t = \mu_{yt} dt + \sigma_{yt} d\omega_t$, where μ_{yt} is as in (16) and

$$\sigma_{yt} = -y_t (1 - y_t) \left[\eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} - \eta_{\mathcal{B}t} \upsilon_{\mathcal{B}t} + Z_t \right]. \tag{A.10}$$

To determine the endogenous υ_{it} terms above, we first note that from (A.3) it follows that the dynamics of $dC_{it}(s)/C_{it}(s)$ are independent of the cohort index s. Therefore, (9) equivalently can be written as

$$d \ln C_{it} - \mathrm{E}_t^i [d \ln C_{it}] = d \ln C_{it}(s) - \mathrm{E}_t^i [d \ln C_{it}(s)] = \upsilon_{it} d\omega_{it},$$

for all s. Since $C_{\mathcal{R}t} = y_t Y_t$ and $C_{\mathcal{B}t} = (1 - y_t) Y_t$, we obtain

$$d \ln C_{\mathcal{R}t} - \mathrm{E}_t [d \ln C_{\mathcal{R}t}] = \left[\sigma + y_t^{-1} \sigma_{yt} \right] d\omega_t, \tag{A.11}$$

and

$$d\ln C_{\mathcal{B}t} - \mathrm{E}_t^{\mathcal{B}}\left[d\ln C_{\mathcal{B}t}\right] = \left[\sigma - (1-y_t)^{-1}\sigma_{yt}\right] d\omega_{\mathcal{B}t}, \tag{A.12}$$

which implies

$$\upsilon_{\mathcal{R}t} = \sigma + y_t^{-1}\sigma_{yt}, \qquad \text{and} \qquad \upsilon_{\mathcal{B}t} = \sigma - (1-y_t)^{-1}\sigma_{yt}. \tag{A.13}$$

Combining (A.10) and (A.13), we obtain (17) and (18).

Finally, to close the model and fully characterize the Markovian equilibria, we need to determine the consumption share of newborn investors, $C_{it}(t)/Y_t$, appearing in (16). From the static budget constraint (A.2), we have

$$\frac{C_{it}(t)}{Y_t} = \frac{\delta}{\pi}\frac{P_t/Y_t}{\mathrm{E}_t\left[\int_t^\infty e^{-\pi(u-t)}\frac{\xi_u}{\xi_t}\frac{C_{iu}(t)}{C_{it}(t)}du\right]}. \tag{A.14}$$

Using (A.4), we obtain the expectation

$$\mathrm{E}_t\left[\int_t^\infty e^{-\pi(u-t)}\frac{\xi_u}{\xi_t}\frac{C_{iu}(t)}{C_{it}(t)}du\right] = \mathrm{E}_t\left[\int_t^\infty e^{-(\rho+\pi)(u-t)}\frac{L_{iu}}{L_{it}}\frac{\gamma_{iu}}{\gamma_{it}}du\right] = \frac{1}{\pi+\rho+\kappa}\left[1+\frac{\kappa}{\pi+\rho}\frac{\overline{\gamma}}{\gamma_{it}}\right], \tag{A.15}$$

where the last equality follows from change-of-measure and risk aversion dynamics (10). In equilibrium, the total wealth must satisfy $P_t = W_{\mathcal{R}t} + W_{\mathcal{B}t}$, where each i type's wealth is given by

$$W_{it} = \int_{-\infty}^t \alpha_i N_t(s)\,\mathrm{E}_t\left[\int_t^\infty e^{-\pi(u-t)}\frac{\xi_u}{\xi_t}C_{iu}(s)\,du\right]ds.$$

The expectation immediately becomes

$$\mathrm{E}_t\left[\int_t^\infty e^{-\pi(u-t)}\frac{\xi_u}{\xi_t}C_{iu}(s)\,du\right] = \frac{1}{\rho+\pi+\kappa}\left[1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{it}}\right]C_{it}(s),$$

which in turn yields each i type's wealth as

$$W_{it} = \frac{1}{\rho+\pi+\kappa}\left[1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{it}}\right]C_{it}. \tag{A.16}$$

As a result, we obtain the aggregate wealth-to-consumption ratio as

$$p_t \equiv \frac{P_t}{Y_t} = \frac{1}{\rho + \pi + \kappa}\left[1 + \frac{\kappa}{\rho + \pi}\frac{\overline{\gamma}}{\Gamma_t}\right], \tag{A.17}$$

where the aggregate risk aversion Γ_t is as in (20). Combining (A.14), (A.15), and (A.17), gives (19). The benchmark economy quantities are obtained by setting $y_t = 1$ in our main economy quantities. □

Proof of Proposition 2. The equilibrium aggregate stock price expression in (21) readily follows from the aggregate wealth-to-consumption ratio (A.17) in the proof of Proposition 1. The benchmark economy stock price is obtained by setting $y_t = 1$ in our main economy quantities, yielding $\Gamma_t = \gamma_{\mathcal{R}t}$. □

Proof of Proposition 3. To determine the equilibrium market risk premium and return volatility, we use the aggregate stock price representation $P_t = p_t Y_t$, where the aggregate wealth-to-consumption (price-dividend) ratio p_t is given by (A.17) in the proof of Proposition 1. We denote the equilibrium dynamics of p_t as

$$\frac{dp_t}{p_t} = \mu_{pt}dt + \sigma_{pt}d\omega_t, \tag{A.18}$$

and obtain aggregate stock price dynamics as

$$\frac{dP_t}{P_t} = \frac{dp_t}{p_t} + \frac{dY_t}{Y_t} + \frac{dp_t}{p_t}\frac{dY_t}{Y_t} = \left[\mu_{pt} + \mu + \sigma\sigma_{pt}\right]dt + \left[\sigma + \sigma_{pt}\right]d\omega_t. \tag{A.19}$$

Therefore, given the postulated stock return dynamics (12), we obtain

$$\mu_{Pt} = \mu_{pt} + \mu + \sigma\sigma_{pt} + \frac{1}{p_t} - \delta, \qquad \text{and} \qquad \sigma_{Pt} = \sigma + \sigma_{pt}. \tag{A.20}$$

Applying Itô's Lemma to the price-dividend ratio in (A.17), we obtain

$$\begin{aligned}
\mu_{pt} = {} & \frac{1}{p_t}\frac{1}{\rho + \pi + \kappa}\frac{\kappa}{\rho + \pi}\left[\left(\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}} - \frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\right)\mu_{yt} + \left(\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t} - \frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t}\right)\sigma_{yt}\right] \\
& - \frac{1}{p_t}\frac{1}{\rho + \pi + \kappa}\frac{\kappa}{\rho + \pi}\left[y_t\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}\frac{1}{\gamma_{\mathcal{R}t}}\kappa\left(\overline{\gamma} - \gamma_{\mathcal{R}t}\right) - \left(1 - y_t\right)\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\frac{1}{\gamma_{\mathcal{B}t}}\kappa\left(\overline{\gamma} - \gamma_{\mathcal{B}t}\right)\right] \\
& + \frac{1}{p_t}\frac{1}{\rho + \pi + \kappa}\frac{\kappa}{\rho + \pi}\left[y_t\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}\eta_{\mathcal{R}t}^2\upsilon_{\mathcal{R}t}^2 + \left(1 - y_t\right)\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\eta_{\mathcal{B}t}^2\upsilon_{\mathcal{B}t}^2 - \left(1 - y_t\right)\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t}Z_t\right],
\end{aligned} \tag{A.21}$$

and

$$\sigma_{pt} = \frac{1}{p_t} \frac{1}{\rho + \pi + \kappa} \frac{\kappa}{\rho + \pi} \left[y_t \frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}} \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} + (1 - y_t) \frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}} \eta_{\mathcal{B}t} \upsilon_{\mathcal{B}t} + \sigma_{yt} \left(\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}} - \frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}} \right) \right]. \tag{A.22}$$

Substituting (A.17) into (A.22) and using (A.20), we obtain the return sensitivity as in (22).

To obtain the market risk premium, we next determine the equilibrium state price density and its dynamics. Given the equilibrium aggregate consumption of type-i investors in (A.7) and the consumption goods market clearing condition $C_{\mathcal{R}t} + C_{\mathcal{B}t} = Y_t$, we have

$$\alpha_{\mathcal{R}} e^{-(\rho+\pi)t} \gamma_{\mathcal{R}t} G_{\mathcal{R}t} \xi_t^{-1} + \alpha_{\mathcal{B}} e^{-(\rho+\pi)t} L_t \gamma_{\mathcal{B}t} G_{\mathcal{B}t} \xi_t^{-1} = Y_t. \tag{A.23}$$

Combining (A.8) and (A.23), we obtain the equilibrium state price density as

$$\xi_t = \alpha_{\mathcal{R}} e^{-(\rho+\pi)t} Y_t^{-1} \gamma_{\mathcal{R}t} G_{\mathcal{R}t} y_t^{-1}. \tag{A.24}$$

Applying Itô's Lemma to the state price density in (A.24), we obtain

$$\begin{aligned} \frac{d\xi_t}{\xi_t} = -\Bigg\{ & \rho + \pi + \mu - \frac{1}{\gamma_{\mathcal{R}t}} \kappa \left(\overline{\gamma} - \gamma_{\mathcal{R}t} \right) - \alpha_{\mathcal{R}} \pi \frac{C_{\mathcal{R}t}(t)}{Y_t} y_t^{-1} + y_t^{-1} \mu_{yt} \\ & - \left(\sigma^2 + y_t^{-2} \sigma_{yt}^2 \right) - \sigma \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} - \sigma y_t^{-1} \sigma_{yt} - \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} y_t^{-1} \sigma_{yt} \Bigg\} dt \\ & - \left[\sigma + \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} + y_t^{-1} \sigma_{yt} \right] d\omega_t, \end{aligned}$$

from its drift term we identify the interest rate as

$$\begin{aligned} r_t = & \rho + \pi + \mu - \frac{1}{\gamma_{\mathcal{R}t}} \kappa \left(\overline{\gamma} - \gamma_{\mathcal{R}t} \right) - \alpha_{\mathcal{R}} \pi \frac{C_{\mathcal{R}t}(t)}{Y_t} y_t^{-1} + y_t^{-1} \mu_{yt} \\ & - \left(\sigma^2 + y_t^{-2} \sigma_{yt}^2 \right) - \sigma \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} - \sigma y_t^{-1} \sigma_{yt} - \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} y_t^{-1} \sigma_{yt}, \end{aligned} \tag{A.25}$$

and from its diffusion term, the market price of aggregate risk as

$$m_t = \sigma + \eta_{\mathcal{R}t} \upsilon_{\mathcal{R}t} + y_t^{-1} \sigma_{yt}. \tag{A.26}$$

Substituting μ_{yt} and σ_{yt} in Proposition 1 into these expressions and rearranging yields

$$r_t = \rho + [\mu + (1-y_t)\,\sigma Z_t] - \left[1 + y_t\frac{\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t}}{\sigma} + (1-y_t)\,\frac{\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t}}{\sigma}\right]\sigma^2$$
$$- \left[y_t\kappa\left(\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}} - 1\right) + (1-y_t)\,\kappa\left(\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}} - 1\right)\right] + \pi\left[1 - \left(\alpha_{\mathcal{R}}\frac{C_{\mathcal{R}t}(t)}{Y_t} + \alpha_{\mathcal{B}}\frac{C_{\mathcal{B}t}(t)}{Y_t}\right)\right], \tag{A.27}$$

$$m_t = \sigma + y_t\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t} + (1-y_t)\,\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t} - (1-y_t)\,Z_t. \tag{A.28}$$

The equilibrium market risk premium is then given by the standard covariance term $-\frac{d\xi_t}{\xi_t}\frac{dP_t}{P_t} = m_t\sigma_{Pt}$, as in (23), which is also equal to $\mu_{Pt} - r_t$ in equilibrium. The benchmark economy quantities are obtained by setting $y_t = 1$ in our main economy quantities. □

Proof of Proposition 4. The equilibrium wealth of investors in (26) follows immediately from (A.16) in the proof of Proposition 1 by substituting $C_{\mathcal{R}t} = y_tY_t$ and $C_{\mathcal{B}t} = (1-y_t)\,Y_t$. The benchmark economy wealth is obtained by setting $y_t = 1$. □

Proof of Proposition 5. To determine investors' portfolios, we first note that, in equilibrium, the financial wealth of an investor of type i born at time s satisfies

$$W_{it}(s) = \mathrm{E}_t\left[\int_t^\infty e^{-\pi(u-t)}\frac{\xi_u}{\xi_t}C_{iu}(s)\,du\right]. \tag{A.29}$$

Substituting (A.3) into (A.29) gives

$$W_{\mathcal{R}t}(s) = \frac{1}{\rho+\pi+\kappa}\left(1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}\right)C_{\mathcal{R}t}(s) = \frac{1}{\rho+\pi+\kappa}\left(1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}\right)\frac{1}{\psi_{\mathcal{R}}}e^{-\rho(t-s)}\gamma_{\mathcal{R}t}\left(\frac{\xi_t}{\xi_s}\right)^{-1}, \tag{A.30}$$

and

$$W_{\mathcal{B}t}(s) = \frac{1}{\rho+\pi+\kappa}\left(1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\right)C_{\mathcal{B}t}(s) = \frac{1}{\rho+\pi+\kappa}\left(1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}\right)\frac{1}{\psi_{\mathcal{B}}}e^{-\rho(t-s)}\gamma_{\mathcal{B}t}\frac{L_t}{L_s}\left(\frac{\xi_t}{\xi_s}\right)^{-1}. \tag{A.31}$$

Applying Itô's Lemma to (A.30) and (A.31), we obtain

$$\frac{dW_{\mathcal{R}t}(s)}{W_{\mathcal{R}t}(s)} = \ldots dt + \left[m_t - \frac{\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t}}{1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}} \right] d\omega_t, \tag{A.32}$$

and

$$\frac{dW_{\mathcal{B}t}(s)}{W_{\mathcal{B}t}(s)} = \ldots dt + \left[m_t + Z_t - \frac{\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t}}{1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}} \right] d\omega_t. \tag{A.33}$$

Matching the diffusion terms in (A.32) and (A.33) with that of the dynamic budget constraint in (14) yields

$$\phi_{\mathcal{R}t}\sigma_{Pt} = m_t - \frac{\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t}}{1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{R}t}}}, \tag{A.34}$$

and

$$\phi_{\mathcal{B}t}\sigma_{Pt} = m_t + Z_t - \frac{\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t}}{1+\frac{\kappa}{\rho+\pi}\frac{\overline{\gamma}}{\gamma_{\mathcal{B}t}}}. \tag{A.35}$$

Dividing both sides of (A.34) and (A.35) by σ_{Pt} and using the identity $\Pi_{Pt} = m_t\sigma_{Pt}$ yields the portfolio formulas in (27). The benchmark economy portfolio fraction is obtained by setting $y_t = 1$, which leads to $\overline{\phi}_{\mathcal{R}t} = 1$. □

Appendix B: Details of Numerical Analysis

Calibration. We estimate the parameters of the aggregate endowment process μ and σ using aggregate consumption data. Specifically, following the literature (e.g., Lettau and Ludvigson (2001) and Bansal and Yaron (2004)), we obtain the quarterly real per-capita consumption series for nondurable goods and services from the Bureau of Economic Analysis (BEA). We then sum these two series to construct the quarterly aggregate consumption time series for the sample period 1990Q1-2023Q4, matching our sample for the market return data. Given the geometric Brownian motion representation of the aggregate endowment process, we set μ and σ such that they are equal to the mean and standard deviation of the aggregate consumption growth rate in our sample, resulting in $\mu = 0.015$ and $\sigma = 0.024$, which are in line with their commonly used values in standard consumption-based asset-pricing models.

Parameter	Symbol	Value
Mean growth rate of aggregate endowment	μ	0.015
Volatility of aggregate endowment	σ	0.024
Population share of rational investors	$\alpha_{\mathcal{R}}$	0.5
Degree of extrapolation	β	0.5
Risk aversion long-run mean	$\overline{\gamma}$	10
Risk aversion lower-bound	λ	6
Risk aversion mean-reversion speed	κ	0.6
Risk aversion sensitivity to shocks	ζ	21.9
Birth/death rate	π	0.03
Depreciation rate	δ	0.02
Subjective time discount rate	ρ	0.02

Table B1. Parameter values. This table reports the parameter values used in our numerical analysis.

For clarity of the analysis, we simply assume that the two investor types each account for half of the total investor population, so $\alpha_{\mathcal{R}} = \alpha_{\mathcal{B}} = 0.5$. Because biased investors form subjective beliefs using extrapolative expectations, we set the extrapolation parameter β to its commonly employed value of 0.5, consistent with its survey-based estimates reported in Barberis et al. (2015) and Cassella and Gulen (2018).

For the risk-aversion parameters, we set its long-run mean to $\overline{\gamma} = 10$, in line with the commonly considered risk aversion values in asset pricing (e.g., Bansal and Yaron (2004)). For the remaining risk-aversion parameters, we choose values so that the model can reasonably match unconditional asset-pricing moments in the data while also satisfying the parameter restriction in (11). Specifically, we set the mean-reversion speed of risk aversion to $\kappa = 0.6$, which implies a shock half-life around 14 months. We set the risk aversion lower bound to $\lambda = 6$, and the sensitivity of the risk aversion to aggregate shocks to close to its upper bound as in Menzly et al. (2004), $\zeta = 21.9$. For the OLG structure parameters, we set the birth-and-death rate to $\pi = 0.03$ and the depreciation rate to $\delta = 0.02$, in line with the literature (e.g., Gârleanu and Panageas (2015) and Gârleanu et al. (2026)). We also choose a plausible subjective time-discount rate of $\rho = 0.02$. The full set of parameter values is summarized in Table B1.

Under these parameter values, the mean annualized real interest rate, market risk premium, and return volatility are 1.28%, 5.25%, and 20.12%, respectively, all close to their historical averages.

Variable	Mean	Std	Min	5%	25%	50%	75%	95%	Max
Z_t	0.002	0.496	-1.583	-0.818	-0.337	0.000	0.337	0.822	1.844
y_t	0.890	0.077	0.097	0.749	0.886	0.916	0.929	0.937	0.957
$\gamma_{\mathcal{R}t}$	9.826	2.364	6.430	7.818	8.422	9.182	10.403	13.915	35.311
$\gamma_{\mathcal{B}t}$	12.129	2.584	6.001	6.580	10.612	12.875	14.076	14.970	20.571
Γ_t	9.865	1.785	6.514	7.767	8.620	9.464	10.653	13.393	27.004

Table B2. Summary statistics of state variables. This table reports the summary statistics for the state variables $(Z_t, \gamma_{\mathcal{R}t}, \gamma_{\mathcal{B}t}, y_t)$ as well as the aggregate risk aversion Γ_t, in our simulated sample.

Simulation. In our model, the equilibrium is characterized by four state variables $(Z_t, \gamma_{\mathcal{R}t}, \gamma_{\mathcal{B}t}, y_t)$, all driven by a common Brownian motion. Because these variables evolve jointly, the effect of any single variable cannot be isolated while holding the others fixed. We therefore rely on numerical simulations to capture their interaction.

To this end, we initialize the system at $(Z_0, \gamma_{\mathcal{R}0}, \gamma_{\mathcal{B}0}, y_0) = (0, \overline{\gamma}, \overline{\gamma}, \alpha_{\mathcal{R}})$.[22] We then simulate the state variables using a standard Euler scheme and the following dynamics under the objective measure $\mathbb{P}$

$$dZ_t = -\beta Z_t dt + \beta d\omega_t, \tag{B.1}$$

$$d\gamma_{\mathcal{R}t} = \kappa\left(\overline{\gamma} - \gamma_{\mathcal{R}t}\right) dt - \gamma_{\mathcal{R}t}\eta_{\mathcal{R}t}\upsilon_{\mathcal{R}t} d\omega_t, \tag{B.2}$$

$$d\gamma_{\mathcal{B}t} = \left[\kappa\left(\overline{\gamma} - \gamma_{\mathcal{B}t}\right) + \gamma_{\mathcal{B}t}\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t} Z_t\right] dt - \gamma_{\mathcal{B}t}\eta_{\mathcal{B}t}\upsilon_{\mathcal{B}t} d\omega_t, \tag{B.3}$$

$$dy_t = \mu_{yt} dt + \sigma_{yt} d\omega_t. \tag{B.4}$$

To improve numerical stability, rather than working directly with the dynamics of unbounded γ_i, we find it helpful to transform to its inverse, $\theta_i \equiv 1/\gamma_i$, which is bounded above by $1/\lambda$. We then apply an exponential Euler discretization to the transformed process, θ_i, which captures risk tolerance (equivalently, the surplus-consumption ratio). To avoid contamination from unrealistically large risk-aversion values, we impose a lower bound of 0.001 on θ_i, corresponding to an upper bound of 1000 for γ_i. We also impose reflecting boundaries for $\theta_i \in (0.001, 1/\lambda)$ and $y \in (0, 1)$ whenever simulated paths cross these bounds. We simulate the model at a daily frequency, using 252 business days per year for 200 years across 5,000 sample paths. We retain one observation every 21 days to construct a monthly sample.

[22] While these values are chosen for convenience, the results are insensitive to the initial conditions, as we discard an initial burn-in period and focus on the simulated sample that approximates the stationary distribution. In addition, we set the initial endowment $Y_0 = 1$ in the simulation for simplicity.

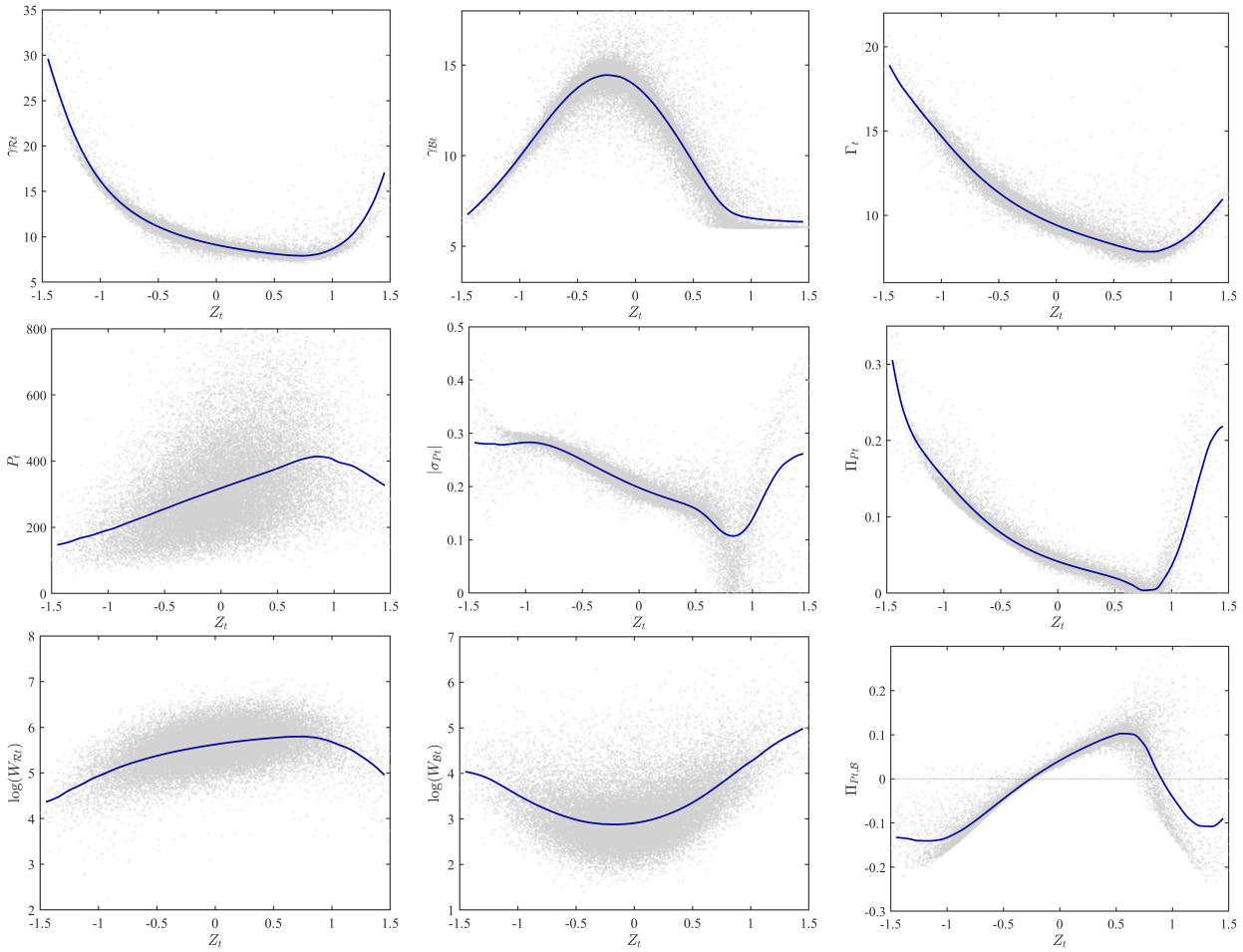


Figure B1. Simulated observations and conditional-bin-mean fitted curves. This figure plots simulated stationary observations from the baseline economy against the state variable Z_t. The gray dots represent simulated observations, and the solid blue curves are fitted through conditional bin means. The panels report, respectively, rational investors' risk aversion $\gamma_{\mathcal{R}t}$, biased investors' risk aversion $\gamma_{\mathcal{B}t}$, aggregate risk aversion Γ_t, aggregate stock price P_t, market return volatility $|\sigma_{Pt}|$, market risk premium Π_{Pt}, rational investors' wealth $\log(W_{\mathcal{R}t})$, biased investors' wealth, $\log(W_{\mathcal{B}t})$, and biased investors' subjective market risk premium $\Pi_{Pt,\mathcal{B}}$.

To approximate the stationary distribution, we discard the first 80% of the simulated observations as a burn-in period and retain the remaining 20%, yielding 40 years of monthly data, comparable to the length of the empirical sample we use. To rule out unrealistically high risk-aversion states, we trim the stationary sample by eliminating the union of states in which $\gamma_{\mathcal{R}t}$ lies in its top 0.25 percentile and states in which $\gamma_{\mathcal{B}t}$ lies in its top 0.25 percentile. This procedure retains 99.5% of the original sample and yields 2,392,986 observations of the state vector $(Z_t, \gamma_{\mathcal{R}t}, \gamma_{\mathcal{B}t}, y_t)$, which provides a sufficiently large sample to study the model's asset pricing implications. Table B2 reports distributional summary statistics for the

simulated state variables and aggregate risk aversion in the stationary equilibrium of our model.

Once we obtain the stationary sample of the state variables, we recover all relevant economic quantities using their closed-form expressions in the model. To illustrate how an economic quantity $\mathcal{Y}$ varies with a conditioning variable $\mathcal{X}$, we use a nonparametric conditional-bin-mean approach. Specifically, for each quantity $\mathcal{Y}$ plotted against $\mathcal{X}$, we partition the simulated support of 20 evenly spaced bins and compute the sample average of $\mathcal{Y}$ within each bin. In the main text, we plot a smooth curve fitted through these conditional bin means. This procedure highlights the average nonlinear relation between $\mathcal{Y}$ and $\mathcal{X}$ while averaging out variation due to the remaining state variables. To show that the conditional-bin-mean fitted curves plotted in the main text capture the nonlinear patterns in the simulated sample, Figure B1 reports the underlying simulated observations, shown as gray dots, together with the fitted curves for the main endogenous quantities. The fitted curves closely trace the central tendency of the simulated observations and therefore provide a concise summary of the model-implied nonlinear relations.

Appendix C: Further Empirical Analysis

This appendix provides additional validation exercises for the empirical analysis in Section 4. We begin by repeating the tests with alternative constructions of the state variable Z_t. We then examine the results obtained under different proxies for conditional return volatility and the risk premium and reassess the findings.

Alternative measures of Z_t. In our baseline empirical analysis, we use aggregate consumption and GDP, both at quarterly frequency, as empirical proxies for the aggregate endowment, and thus for our business cycle measure Z_t. Here, we consider two additional time series commonly used to capture aggregate economic conditions: the industrial production index (INDPRO) from FRED and the real earnings series from Robert Shiller's dataset, both available at monthly frequency. We treat these as alternative proxies for Y_t and apply the same procedure as in Section 4 to construct normalized state variables. This yields an INDPRO-based and a real-earnings-based Z_t series. We use the same

conditional risk-premium and return-volatility series over the 1990–2023 sample period. Using these series, we run the same set of regressions in Section 4 and obtain Table C1, which reports the results for the full-sample linear regression as well as the continuous piecewise linear regressions with cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$.

As shown in Table C1, the results closely mirror our baseline findings: the positive relation between the asset pricing quantities and Z_t emerges within broad expansions and becomes more economically significant in strong expansions.

Since Z_t in the model also captures belief distortions, we further consider survey-based measures of consumer beliefs from the University of Michigan Surveys of Consumers. Specifically, we use the headline Index of Consumer Sentiment (ICS), which captures broad household assessments of current and expected economic conditions, and the Index of Consumer Expectations (ICE), which focuses more directly on beliefs about future economic conditions. We normalize each series by subtracting its sample mean and dividing by its sample standard deviation, and treat the resulting standardized series as alternative sentiment- and expectation-based proxies for Z_t. We use the same conditional risk-premium and return-volatility series over the 1990–2023 sample period. Since the sentiment and expectation measures are monthly, we match them with the asset-pricing quantities measured on the last trading day of each month. Using these series, we run the same set of regressions as in Section 4 and obtain Table C2, which reports the results for the full-sample linear regression as well as the continuous piecewise linear regressions with cutoffs $c = 0$ and $c = 1$.[23]

As shown in Table C2, the results are robust to using survey-based proxies for Z_t. For both sentiment and expectations, the positive relation between the asset-pricing quantities and Z_t emerges in expansions and becomes substantially stronger in high-sentiment or high-expectation states, consistent with our model implications.

[23] Because ICS and ICE are standardized, positive values indicate above-average optimism and negative values indicate below-average optimism, in line with the definition of belief distortions. The cutoff $c = 1$ therefore corresponds to one standard deviation above the mean, consistent with our baseline analysis.

Alternative measures of volatility. In our baseline empirical analysis, we use the conditional return volatility estimate from Schreindorfer and Sichert (2025). Here, we consider two alternative volatility measures. Our first volatility measure is constructed from the realized-variance series available on Hao Zhou's website, which extends the high-frequency (5-minute) S&P 500 index return data based measures used in Bollerslev et al. (2009). From this series, we obtain a monthly return-volatility measure, which we refer to as Vol_BTZ. The second volatility measure is based on daily S&P 500 index returns. Using daily returns from CRSP/WRDS, denoted by r_d, we compute for each month m the realized variance as the sum of squared daily log returns, $RV_m = \sum_{d \in m} r_d^2$. We then obtain the annualized monthly volatility by $\sqrt{12 \times RV_m}$. We refer to this series as Vol_Our. Because the conditional return volatility used in the main text is the time-t conditional volatility of one-month-ahead returns, the natural volatility counterpart is the volatility realized over the following month. We therefore estimate the predictive piecewise linear regression

$$\mathcal{Y}_{t+1} = \alpha + \beta_1 Z_t + \beta_2 \left(Z_t - c\right)^+ + \varepsilon_{t+1}, \tag{C.1}$$

and the predictive full-sample linear regression

$$\mathcal{Y}_{t+1} = \overline{\alpha} + \overline{\beta} Z_t + \epsilon_{t+1}, \tag{C.2}$$

where the independent variable is either the consumption-based or real GDP-based Z_t in the current quarter, and the dependent variable $\mathcal{Y}_{t+1}$ is realized return volatility in the first month of the next quarter. The matched regression sample runs from 1990Q1 to 2024Q3. We report the regression results for the predictive full-sample linear regression (C.2) and the predictive piecewise linear regressions (C.1) with cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$ in Table C3.

Although forward realized return volatility is a noisy ex-post proxy for conditional return volatility, the results display the same pattern as in the main text: under both volatility measures, one-month-ahead realized volatility rises with Z_t in broad expansions, and the positive relation becomes much stronger in strong expansions.

Alternative measures of risk premium. In our baseline empirical analysis, we use the conditional risk premium estimate from Schreindorfer and Sichert (2025). In this robustness check, we replace it with two alternative measures. The first measure is option-implied lower bound for the market equity premium as in Martin (2017), who shows that this lower bound is typically tight. We obtain this lower-bound series from Amit Goyal's website. Because the lower bound is already forward-looking, we run contemporaneous regressions of the lower bound on the current state variable Z_t, using the same full-sample linear and continuous piecewise linear specifications as in the main text. Specifically, we match the current-quarter Z_t to the lower-bound observation in the last month of the same quarter. The lower-bound sample runs from 1996Q1 to 2023Q3. The second measure is the annualized one-month realized excess return, computed as the monthly S&P 500 return minus the one-month risk-free rate. The monthly stock return is from CRSP/WRDS, and the risk-free rate is from Kenneth French's dataset. To match the one-month-ahead horizon of the conditional risk premium in the main text, we run predictive regressions of next-month realized excess returns on the current state variable Z_t as in (C.1) and (C.2). Specifically, we match the current-quarter Z_t to the realized excess return in the first month of the next quarter. Since return and risk-free-rate data are available over a much longer sample, we run the realized-return regressions from 1960Q1 to 2025Q3. The regression results are reported in Table C4.

Comparing the right-side slope $\beta_1 + \beta_2$ across the cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$, the positive relation between proxied risk premia and Z_t becomes much stronger in strong expansions, even under the noisy realized-excess-return proxy, as in our baseline analysis.

	INDPRO-based Z_t		Real earnings-based Z_t	
	Volatility	Risk premium	Volatility	Risk premium
Panel A: full-sample linear regression				
$\overline{\alpha}$	0.142^{***}	0.087^{***}	0.142^{***}	0.086^{***}
	(0.006)	(0.004)	(0.005)	(0.004)
$\overline{\beta}$	-0.021	-0.019^{**}	-0.026^{**}	-0.013^{*}
	(0.014)	(0.010)	(0.010)	(0.008)
Observations	408	408	408	408
R^2	0.045	0.067	0.128	0.060
Panel B: piecewise regression with cutoff $c = 0$				
α	0.122^{***}	0.071^{***}	0.126^{***}	0.073^{***}
	(0.007)	(0.005)	(0.006)	(0.005)
β_1	-0.065^{***}	-0.053^{***}	-0.053^{***}	-0.035^{***}
	(0.018)	(0.011)	(0.012)	(0.008)
$\beta_1 + \beta_2$	0.033^{**}	0.022^{*}	0.013	0.017^{***}
	(0.013)	(0.012)	(0.009)	(0.006)
Obs.$(Z_t \leq c)$/Obs.$(Z_t > c)$	203/205	203/205	186/222	186/222
R^2	0.170	0.195	0.246	0.192
Panel C: piecewise regression with cutoff $c = \overline{Z} + 1 \times \text{std}$				
α	0.131^{***}	0.077^{***}	0.136^{***}	0.081^{***}
	(0.005)	(0.003)	(0.005)	(0.004)
β_1	-0.054^{***}	-0.047^{***}	-0.044^{***}	-0.028^{***}
	(0.013)	(0.008)	(0.009)	(0.006)
$\beta_1 + \beta_2$	0.180^{***}	0.150^{***}	0.059^{***}	0.056^{***}
	(0.034)	(0.025)	(0.013)	(0.010)
Obs.$(Z_t \leq c)$/Obs.$(Z_t > c)$	340/68	340/68	376/32	376/32
R^2	0.221	0.286	0.260	0.213

Table C1. Regression results using alternative Z_t measures based on INDPRO and earnings. This table reports regressions of conditional return volatility and the conditional risk premium on alternative state variables Z_t. The first two columns use the INDPRO-based Z_t, while the last two columns use the real-earnings-based Z_t. Both dependent variables are from Schreindorfer and Sichert (2025) and are measured on the last trading day of each month, using the matched monthly sample from 1990 to 2023. Panel A reports the full-sample linear regression (25). Panels B and C report the continuous piecewise linear regression (24) with cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$, respectively. In the piecewise regressions, the table reports α, the slope β_1 for $Z_t \leq c$, and the high-Z_t slope $\beta_1 + \beta_2$ for $Z_t > c$. Newey-West standard errors are reported in parentheses, using five monthly lags. Statistical significance is denoted by *, **, ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

	Michigan ICS		Michigan ICE	
	Volatility	Risk premium	Volatility	Risk premium
Panel A: full-sample linear regression				
$\overline{\alpha}$	0.142^{***}	0.086^{***}	0.142^{***}	0.086^{***}
	(0.006)	(0.004)	(0.006)	(0.004)
$\overline{\beta}$	-0.014^{*}	-0.011^{**}	-0.007	-0.007
	(0.007)	(0.005)	(0.007)	(0.005)
Observations	408	408	408	408
R^2	0.058	0.064	0.015	0.024
Panel B: piecewise regression with cutoff $c = 0$				
α	0.109^{***}	0.065^{***}	0.107^{***}	0.064^{***}
	(0.009)	(0.006)	(0.008)	(0.006)
β_1	-0.049^{**}	-0.033^{***}	-0.049^{***}	-0.033^{***}
	(0.013)	(0.009)	(0.012)	(0.008)
$\beta_1 + \beta_2$	0.032^{***}	0.018^{**}	0.037^{***}	0.021^{***}
	(0.009)	(0.008)	(0.007)	(0.007)
Obs.$(Z_t \leq c)$/Obs.$(Z_t > c)$	181/227	181/227	193/215	193/215
R^2	0.200	0.167	0.194	0.147
Panel C: piecewise regression with cutoff $c = 1$				
α	0.133^{***}	0.081^{***}	0.133^{***}	0.081^{***}
	(0.005)	(0.004)	(0.005)	(0.004)
β_1	-0.028^{***}	-0.020^{***}	-0.023^{***}	-0.017^{***}
	(0.007)	(0.005)	(0.008)	(0.005)
$\beta_1 + \beta_2$	0.121^{***}	0.074^{***}	0.085^{***}	0.052^{***}
	(0.017)	(0.018)	(0.014)	(0.015)
Obs.$(Z_t \leq c)$/Obs.$(Z_t > c)$	353/55	353/55	352/56	352/56
R^2	0.209	0.169	0.147	0.119

Table C2. Regression results using alternative Z_t measures based on sentiment and expectations. This table reports regressions of conditional return volatility and conditional risk premium on survey-based proxies for Z_t. The first two columns use the normalized Michigan Index of Consumer Sentiment (ICS), and the last two columns use the normalized Michigan Index of Consumer Expectations (ICE). The dependent variables are the conditional return volatility and conditional risk premium from Schreindorfer and Sichert (2025), measured on the last trading day of each month over the 1990–2023 sample. Panel A reports the full-sample linear regression. Panels B and C report the continuous piecewise linear regressions with cutoffs $c = 0$ and $c = 1$, respectively. In the piecewise regressions, the table reports α, the slope β_1 for $Z_t \leq c$, and the high-Z_t slope $\beta_1 + \beta_2$ for $Z_t > c$. Newey-West standard errors with four monthly lags are reported in parentheses. Statistical significance is denoted by *, **, ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

	Consumption-based Z_t		GDP-based Z_t	
	Vol_BTZ	Vol_Our	Vol_BTZ	Vol_Our
Panel A: full-sample linear regression				
$\overline{\alpha}$	0.141^{***}	0.156^{***}	0.141^{***}	0.156^{***}
	(0.011)	(0.012)	(0.011)	(0.012)
$\overline{\beta}$	-0.005	0.006	-0.021	-0.010
	(0.024)	(0.026)	(0.025)	(0.030)
Observations	139	139	139	139
R^2	0.005	0.001	0.011	0.002
Panel B: piecewise regression with cutoff $c = 0$				
α	0.122^{***}	0.139^{***}	0.123^{***}	0.133^{***}
	(0.014)	(0.016)	(0.014)	(0.016)
β_1	-0.057^{**}	-0.043	-0.066^{**}	-0.069^{*}
	(0.027)	(0.033)	(0.029)	(0.036)
$\beta_1 + \beta_2$	0.083^{**}	0.090^{*}	0.048	0.078^{**}
	(0.038)	(0.047)	(0.031)	(0.036)
Obs.$(Z_t \leq c)$/Obs.$(Z_t > c)$	62/77	62/77	61/78	61/78
R^2	0.045	0.034	0.050	0.055
Panel C: piecewise regression with cutoff $c = \overline{Z} + 1 \times \text{std}$				
α	0.134^{***}	0.150^{***}	0.133^{***}	0.146^{***}
	(0.011)	(0.012)	(0.010)	(0.011)
β_1	-0.033	-0.018	-0.051^{**}	-0.049^{*}
	(0.023)	(0.027)	(0.025)	(0.028)
$\beta_1 + \beta_2$	0.233^{***}	0.212^{*}	0.215^{***}	0.291^{***}
	(0.086)	(0.118)	(0.046)	(0.056)
Obs.$(Z_t \leq c)$/Obs.$(Z_t > c)$	121/18	121/18	123/16	123/16
R^2	0.041	0.026	0.067	0.077

Table C3. Regression results using alternative volatility measures. This table reports predictive regressions of realized return volatility on the state variable Z_t. The first two columns use the consumption-based Z_t, while the last two columns use the real GDP-based Z_t. "Zhou Vol" denotes the monthly annualized realized volatility constructed from Hao Zhou's realized variance series, and "Constructed Vol" denotes the monthly annualized realized volatility constructed from daily S&P 500 index returns. The matched regression sample runs from 1990Q1 to 2024Q3. Panel A reports the full-sample linear regression. Panels B and C report the continuous piecewise linear regression with cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$, respectively. In the piecewise regressions, the table reports α, the slope β_1 for $Z_t \leq c$, and the high-Z_t slope $\beta_1 + \beta_2$ for $Z_t > c$. Newey-West standard errors with four quarterly lags are reported in parentheses. Statistical significance is denoted by *, **, ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

	Consumption-based Z_t		GDP-based Z_t	
	Lower bound	Excess return	Lower bound	Excess return
Panel A: full-sample linear regression				
$\overline{\alpha}$	0.042^{***}	0.090^{***}	0.042^{***}	0.093^{***}
	(0.005)	(0.030)	(0.005)	(0.030)
$\overline{\beta}$	-0.013	-0.130^{**}	-0.022^{*}	-0.116^{**}
	(0.011)	(0.056)	(0.012)	(0.057)
Observations	111	263	111	263
R^2	0.018	0.013	0.063	0.013
Panel B: piecewise regression with cutoff $c = 0$				
α	0.035^{***}	0.065	0.032^{***}	0.005
	(0.006)	(0.041)	(0.006)	(0.043)
β_1	-0.031^{**}	-0.184^{***}	-0.045^{***}	-0.317^{***}
	(0.013)	(0.066)	(0.014)	(0.096)
$\beta_1 + \beta_2$	0.018	-0.052	0.016	0.094
	(0.017)	(0.121)	(0.016)	(0.081)
Obs.$(Z_t \le c)$/Obs.$(Z_t > c)$	49/62	133/130	49/62	135/128
R^2	0.046	0.015	0.117	0.030
Panel C: piecewise regression with cutoff $c = \overline{Z} + 1 \times \text{std}$				
α	0.040^{***}	0.083^{***}	0.038^{***}	0.071^{**}
	(0.005)	(0.031)	(0.004)	(0.031)
β_1	-0.020^{*}	-0.152^{**}	-0.035^{***}	-0.199^{***}
	(0.012)	(0.061)	(0.011)	(0.064)
$\beta_1 + \beta_2$	0.056	0.123	0.098^{***}	0.404^{**}
	(0.040)	(0.321)	(0.035)	(0.169)
Obs.$(Z_t \le c)$/Obs.$(Z_t > c)$	94/17	221/42	96/15	224/39
R^2	0.033	0.015	0.120	0.027

Table C4. Regression results using alternative risk premium measures. This table reports regressions of alternative risk-premium measures on the state variable Z_t. The first two columns use the consumption-based Z_t, while the last two columns use the real GDP-based Z_t. "Lower bound" denotes Martin's one-month-ahead option-implied lower bound, and "Excess return" denotes the annualized one-month realized S&P 500 excess return. The lower-bound regressions are contemporaneous and use the 1996Q1–2023Q3 sample; the excess-return regressions are predictive and use the 1960Q1–2025Q3 sample. Panel A reports the full-sample linear regression. Panels B and C report the continuous piecewise linear regression with cutoffs $c = 0$ and $c = \overline{Z} + 1 \times \text{std}$, respectively. In the piecewise regressions, the table reports α, the slope β_1 for $Z_t \le c$, and the high-Z_t slope $\beta_1 + \beta_2$ for $Z_t > c$. Newey-West standard errors with four quarterly lags are reported in parentheses. Statistical significance is denoted by *, **, ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

References

Abel, Andrew B., 1990, Asset prices under habit formation and catching up with the joneses, *American Economic Review* 80, 38–42.

Abel, Andrew B, 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.

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

Afrouzi, Hassan, Spencer Y. Kwon, Augustin Landier, Yueran Ma, and David Thesmar, 2023, Overreaction in expectations: Evidence and theory, *Quarterly Journal of Economics* 138, 1713–1764.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Clara Vega, 2007, Real-time price discovery in global stock, bond and foreign exchange markets, *Journal of international Economics* 73, 251–277.

Atmaz, Adem, 2022, Stock return extrapolation, option prices, and variance risk premium, *Review of Financial Studies* 35, 1348–1393.

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

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

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, Extrapolation and bubbles, *Journal of Financial Economics* 129, 203–227.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

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

Blanchard, Olivier J., 1985, Debt, deficits, and finite horizons, *Journal of Political Economy* 93, 223–247.

Boldrin, Michele, Lawrence J. Christiano, and Jonas D. M. Fisher, 2001, Habit persistence, asset returns, and the business cycle, *American Economic Review* 91, 149–166.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Boyd, John H., Jian Hu, and Ravi Jagannathan, 2005, The stock market's reaction to unemployment news: Why bad news is usually good for stocks, *Journal of Finance* 60, 649–672.

Bricker, Jesse, Brian Bucks, Arthur Kennickell, Traci Mach, and Kevin Moore, 2014, Drowning or weathering the storm? changes in family finances from 2007 to 2009, in *Measuring wealth and financial intermediation and their links to the real economy*, 323–348 (University of Chicago Press).

Brunnermeier, Markus K., and Stefan Nagel, 2008, Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals' asset allocation, *American Economic Review* 98, 713–736.

Brunnermeier, Markus K., and Yuliy Sannikov, 2014, A macroeconomic model with a financial sector, *American Economic Review* 104, 379–421.

Calvet, Laurent E., and Paolo Sodini, 2014, Twin picks: Disentangling the determinants of risk-taking in household portfolios, *Journal of Finance* 69, 867–906.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Cassella, Stefano, and Huseyin Gulen, 2018, Extrapolation bias and the predictability of stock returns by price-scaled variables, *Review of Financial Studies* 31, 4345–4397.

Chabakauri, Georgy, 2013, Dynamic equilibrium with two stocks, heterogeneous investors, and portfolio constraints, *Review of Financial Studies* 26, 3104–3141.

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

Choi, James J., and Thomas M. Mertens, 2013, Extrapolative expectations and the equity premium, Working paper, Yale University.

Cohn, Alain, Jan Engelmann, Ernst Fehr, and Michel André Maréchal, 2015, Evidence for countercyclical risk aversion: An experiment with financial professionals, *American Economic Review* 105, 860–885.

Constantinides, George M., 1990, Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519–543.

Cox, John C., and Chi-fu Huang, 1989, Optimal consumption and portfolio policies when asset prices follow a diffusion process, *Journal of Economic Theory* 49, 33–83.

Detemple, Jerome B., and Fernando Zapatero, 1991, Asset prices in an exchange economy with habit formation, *Econometrica* 59, 1633–1657.

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen, 2017, Asset prices and portfolio choice with learning from experience, *Review of Economic Studies* 85, 1752–1780.

Ferson, Wayne E., and George M. Constantinides, 1991, Habit persistence and durability in aggregate consumption: Empirical tests, *Journal of Financial Economics* 29, 199–240.

Gallmeyer, Michael, and Burton Hollifield, 2008, An examination of heterogeneous beliefs with a short-sale constraint in a dynamic economy, *Review of Finance* 12, 323–364.

Gârleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Gârleanu, Nicolae, Stavros Panageas, and Geoffery Zheng, 2026, A long and a short leg make for a wobbly equilibrium, *American Economic Review* 116, 1234–1273.

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

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2018, Time varying risk aversion, *Journal of Financial Economics* 128, 403–421.

He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 732–770.

Heyerdahl-Larsen, Christian, 2014, Asset prices and real exchange rates with deep habits, *Review of Financial Studies* 27, 3280–3317.

Hirshleifer, David, Jun Li, and Jianfeng Yu, 2015, Asset pricing in production economies with extrapolative expectations, *Journal of Monetary Economics* 76, 87–106.

Jermann, Urban J., 1998, Asset pricing in production economies, *Journal of Monetary Economics* 41, 257–275.

Jin, Lawrence J., and Pengfei Sui, 2022, Asset pricing with return extrapolation, *Journal of Financial Economics* 145, 273–295.

Karatzas, Ioannis, John P. Lehoczky, and Steven E. Shreve, 1987, Optimal portfolio and consumption decisions for a small investor on a finite horizon, *SIAM Journal on Control and Optimization* 25, 1557–1586.

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

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

McQueen, Grant, and V. Vance Roley, 1993, Stock prices, news, and business conditions, *Review of Financial Studies* 6, 683–707.

Meeuwis, Maarten, 2020, Wealth fluctuations and risk preferences: Evidence from US investor portfolios, Working paper, Washington University in St. Louis.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112, 1–47.

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

Paravisini, Daniel, Veronica Rappoport, and Enrichetta Ravina, 2017, Risk aversion and wealth: Evidence from person-to-person lending portfolios, *Management Science* 63, 279–297.

Santos, Tano, and Pietro Veronesi, 2010, Habit formation, the cross section of stock returns and the cash-flow risk puzzle, *Journal of Financial Economics* 98, 385–413.

Santos, Tano, and Pietro Veronesi, 2022, Leverage, *Journal of Financial Economics* 145, 362–386.

Schreindorfer, David, and Tobias Sichert, 2025, Conditional risk and the pricing kernel, *Journal of Financial Economics* 171, 104106.

Simsek, Alp, 2021, The macroeconomics of financial speculation, *Annual Review of Economics* 13, 335–369.

Sundaresan, Suresh M., 1989, Intertemporally dependent preferences and the volatility of consumption and wealth, *Review of Financial Studies* 2, 73–89.

Verdelhan, Adrien, 2010, A habit-based explanation of the exchange rate risk premium, *Journal of Finance* 65, 123–146.

Veronesi, Pietro, 2019, Heterogeneous households under uncertainty, Working paper, National Bureau of Economic Research (No. w25448).

Vissing-Jorgensen, Annette, 2003, Perspectives on behavioral finance: Does "irrationality" disappear with wealth: Evidence from expectations and actions, *NBER Macroeconomics Annual* 18, 139–194.

Wachter, Jessica A., 2006, A consumption-based model of the term structure of interest rates, *Journal*

of Financial Economics 79, 365–399.

Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.

Wachter, Jessica A., and Motohiro Yogo, 2010, Why do household portfolio shares rise in wealth?, *Review of Financial Studies* 23, 3929–3965.

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