Real Term Premia Explained by Stochastic Volatility*

Errikos Melissinos†

15 September 2024

Latest Version

Abstract

Real term premia of risk-free bonds exhibit substantial variation over time. This paper shows that stochastic volatility can give rise to such behaviour, with the required amount of volatility in consumption or wealth being similar to the amount of volatility in the stock market. This explanation is consistent with both the intermediary asset pricing approach and with supply and demand driven explanations, in both of which a small group of investors play a special role in the pricing of securities. Furthermore, under standard preferences and without time-varying risk aversion such stochastic volatility is also necessary to explain real term premia. The paper analyses models with both time-separable and recursive preferences. The latter model variations are solved via a novel perturbation method with respect to the parameter for intertemporal elasticity of substitution.

JEL: C65, E43, G12

Keywords: term premia, stochastic volatility, habit, long-run risk, limited arbitrage, high consumption volatility, high wealth volatility, recursive utility, solution methods

*For helpful comments I thank Mariia Bondar, Julian Detemple, Fabio Girardi, Refet Gürkaynak, Michael Haliassos, Florian Heider, Alexander Hillert, Ruggero Jappeli, Martin Kliem, Holger Kraft, Nora Lamersdorf, Mortiz Lenel, Martin Lettau, Christoph Meinedring, Alexander Meyer-Gohde, Emanuel Mönch, Dmitry Mukhin, Christian Mücke, Loriana Pelizzon, Christian Schlag, Paul Schneider, Nicolas Syrichas, Farah Tohme, Henning Weber, Volker Wieland, the participants of the SAFE Brown Bag Seminar and the participants of the first Bonn-Frankfur t-Mannheim PhD Conference.

†Goethe University Frankfurt & Leibniz Institute for Financial Research SAFE e.V.
website: https://errikos-melissinos.com

1 Introduction

Risk-free bonds lie at the heart of asset pricing theory, as they form the basis for the pricing of all securities. Yet theory still struggles to explain the prices of long-term bonds that we observe in the data. In particular, long-term bonds have been known to violate the expectations hypothesis (EH) (e.g. Fama and Bliss 1987; Cochrane and Piazzesi 2005). This implies that the corresponding

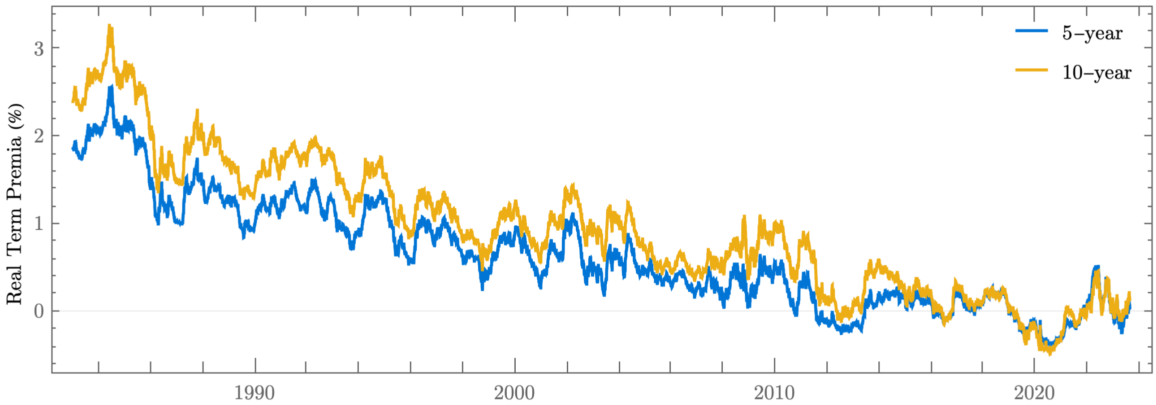


Figure 1: **Time series of real term premia for the US**
Data is taken from d' Amico *et al.* (2018), who decomposed nominal yields into risk-neutral real yields (expected short-term rates averaged over the corresponding period), real term premia, expected inflation, inflation premia and liquidity premia.
Data Source: https://www.federalreserve.gov/econres/notes/feds-notes/tips-from-tips-update-and-discussions-20190521.html

term premia, which reflect the excess expected return of long-term bonds, are time-varying.[1] In addition, more recent literature has also estimated real term premia directly, showing that they are mostly positive and significantly time-varying (Abrahams, Adrian, Crump, Moench and Yu 2016; d' Amico *et al.* 2018; Pflueger and Viceira 2016). Estimates from d' Amico *et al.* (2018) are shown

[1]Real term premia are formally defined as the expected difference in annualised log return from holding long-term bonds compared to short-term bonds over the same period (see also A for mathematical definitions). So, for instance, the ten-year real term premium is the difference in expected annualised log return of a ten-year inflation-adjusted bond compared to the expected annualised return from rolling over instantaneous debt over ten years. In the literature there are two separate formulations of the expectations hypothesis. The strong version suggests that term premia should be equal to zero for all maturities. The weak version suggests that term premia for each maturity are constant over time. The main text refers to the weak version of the expectations hypothesis, which is more commonly analysed. While the first tests of the expectations hypothesis focused on nominal bonds, the focus of this paper is real term premia.

in Figure 1.[2] Overall, models struggle to generate these main features. This is referred to as the bond premium puzzle (Backus, Gregory and Zin 1989).[3] The source of the puzzle is that consumption-based mechanisms typically generate small, negative, and often constant term premia, namely the exact opposite of what we see in the data. This is due to bond prices typically being counter-cyclical in these models, implying negative term premia, which are in any case small in absolute value and practically constant, given that consumption risk is also small and varies little with the business cycle.

After performing a thorough analysis of several model variations, this paper shows that it is possible to explain real term premia by using models with stochastic volatility. The paper focuses on consumption volatility, which can also be thought of as wealth volatility. Crucially, it is shown that the level of volatility needs to be much higher compared to the level of volatility that we see in aggregate consumption data. Indeed, the level of consumption volatility that generates the required level of variation in real term premia, is similar to the level of volatility in the stock market. This suggests that the marginal investors in the bond market are a small group of investors whose wealth is exposed to roughly the amount of volatility as in the stock market. This view is consistent with the intermediary asset-pricing paradigm (He and Krishnamurthy 2013), in which prices are primarily set by intermediaries; and it is also consistent with Vayanos and Vila (2021), who suggested that the term structure of interest rates is determined by arbitrageurs who interact with preferred habitat investors.

In the literature, other models have been able to generate the main features of real term premia by assuming time-varying risk aversion. For instance, Wachter (2006) showed this in a model with an external habit following Campbell and Cochrane (1999). However, when assuming constant risk aversion in a model with a single steady-state-reverting state variable, stochastic volatility becomes not only sufficient but also *necessary* to generate positive and time-varying term premia, as observed in the data. This result holds both with time-separable utility (TSU) and with recursive utility (RU), and it follows from an exhaustive analysis of different model variations and parametrisations. In fact, stochastic volatility is an important ingredient that is also implicitly used in Kekre, Lenel and Mainardi (2022) and Schneider (2022), in order to generate realistic term structures of in-

[2]In these estimates a long-term downward trend stands out (this can be seen even clearer in the longer time-series in Figure E.2). This paper uses a steady-state-reverting state variable. Therefore, the long-term trend is not analysed.

[3]The bond premium puzzle can also refer to nominal term premia. Nominal term premia have the same definition as real term premia with the underlying bonds not being inflation-adjusted.

terest rates, even though these papers' primary focus is not stochastic volatility itself.[4]

Furthermore, the paper provides further contributions. Firstly, it provides explicit values of term premia as a function of the state of the economy for a large range of model variations. This is useful, because consumption-based models in the literature often focus on nominal term premia, and even when they focus on real term premia, explicit state-dependent term premia are rarely displayed. Secondly, it contributes a novel perturbation method to easily and robustly solve models with RU. My perturbation method builds on the approach of Tsai and Wachter (2018). While they used an approximation to the value function that is constant in terms of the intertemporal elasticity of substitution (IES), and analytically correct only for IES equal to 1, the full perturbation series in terms of the IES is derived. This provides a global approximation in terms of the state variable of the economy that allows the easy solution of the model for most values of the IES that are economically interesting. It is also the first perturbation method in terms of the IES within RU models. This method is also explained in further detail in Melissinos (2023).

The rest of the paper is organised as follows: Section 2 contains more information regarding the literature on the bond premium puzzle is provided. Section 3 discusses interest rates in the data. Section 4 presents the setup that will allow to price bonds in the context of TSU and RU. This includes the outline of the novel perturbation method. Section 5 shows and comments on the results for term premia. Finally, section 6 concludes.

2 Literature on the Bond Premium Puzzle

While the paper analyses real term premia, the bond premium puzzle, which refers to the difficulty of standard models to generate positive, large, and time-varying term premia, originally referred to nominal bonds.[5] One of the first papers to address this was Backus *et al.* (1989). Utilising a consumption-based asset-pricing model of an endowment economy, they discovered the model's inability to yield significant positive term premia. Subsequent studies by Donaldson, Johnsen and Mehra (1990) and Den Haan (1995) further indicated that standard real business cycle models also could not resolve the puzzle. Rudebusch and Swanson (2008) incorporated an external habit into DSGE models but found that the bond premium

[4] schneider2022risk also uses effective time-varying risk aversion.

[5] Rudebusch and Swanson (2008) also offered a good summary of this extensive literature.

puzzle remains. Specifically, including a habit with non-flexible working hours can generate positive term premia, but at the cost of inducing volatile wages, prices and short-term interest rates. Duffee (2013) showed that basic properties of nominal yields cannot be explained macroeconomically, at least according to standard asset-pricing models. Also in a more generic contribution, Duffee (2002) shed light on the challenges of fitting both interest rate and term premium dynamics within affine models.

Next, a series of papers provided explanations that focused on nominal term premia, and not on real term premia. Notably, Piazzesi and Schneider (2006) showed that parameter uncertainty in a model where inflation brings bad news about future consumption growth can produce positive nominal term premia.[6] Gabaix (2012) and Tsai (2015), following Rietz (1988) and Barro (2006), showed that positive nominal term premia can be explained, if inflation is on average high during consumption disasters. Bansal and Shaliastovich (2013), following Bansal and Yaron (2004), demonstrated that the risk premium of a nominal bond can be positive in a model with long-run risk, as long as inflation is correlated with consumption trend. Rudebusch and Swanson (2012) used a similar model within a DSGE framework, which has real and nominal long-term risks, and they show that positive nominal term premia are generated; nevertheless real term premia are again negative in this model. Gomez-Cram and Yaron (2021) also used a model following Bansal and Yaron (2004), but they focused on explaining nominal term premia, using an inflation channel, while claiming that the apparent under-performance of their model concerning real term premia should be expected due to liquidity premia in the TIPS market.

Alternatively, some articles also consider real term premia. For instance, Katagiri (2022) explored a model with monetary policy, in which consumption changes can be negatively correlated with consumption trends, and risk aversion is very high. As a result, term premia can be positive, but the premia variability is not examined. Ellison and Tischbirek (2021) went beyond standard rational expectations models by using a beauty contest mechanism as introduced by Angeletos, Collard and Dellas (2018), in which agents anticipate the expectations of other agents; their model generates positive term premia.

Using a similar approach to the current paper, some articles tackle the problem by deviating from the representative agent model. Vayanos and Vila (2021) sug-

[6] Collin-Dufresne, Johannes and Lochstoer (2016) introduced a model with Bayesian learning of parameters. However, this model does not emphasise bond term premia and it generates *negative* term premia.

gested that term premia are generated by arbitrageurs interacting with so-called preferred habitat investors, namely investors that tend to hold specific maturities of bonds. Kekre *et al.* (2022) built on Vayanos and Vila (2021), and showed that the characteristics of the arbitrageur portfolio can have important implications for the sign of term premia. Jappelli, Subrahmanyam and Pelizzon (2023) also built on Vayanos and Vila (2021) by integrating the repo market in their analysis. Schneider (2022) showed that positive term premia can arise in models with heterogeneous agents exhibiting different attitudes towards risk and different preferences to substituting consumption through time. Finally, returning to models with a representative agent, Wachter (2006) showed that term premia can be positive and time-varying, within a model with an external habit following Campbell and Cochrane (1999). Kliem and Meyer-Gohde (2022) used the same mechanism within a DSGE model, and they found positive term premia. Hsu, Li and Palomino (2021) also used this mechanism within a DSGE model, and they verified that a habit element is key in generating positive and time-varying term premia. Campbell, Pflueger and Viceira (2020) also used a habit model to explain the time-variability of term premia. More generally, a model with external habit can be classified as a model with time-varying effective risk aversion, and within this class of models, Lettau and Wachter (2011) showed that positive and time-varying term premia can be obtained, and Bekaert, Engstrom and Grenadier (2010) showed that time-varying term premia can be obtained. These papers all use time-varying risk aversion, which is to my knowledge the only mechanism in the literature that achieves positive and time-varying term premia within a rational representative agent model.[7]

3 Real Rates in the Data

3.1 TIPS as real rates

The first challenge regarding real rates is that they are not directly observable from standard bonds. The real interest rate is the yield of a nominal bond whose payoff is adjusted for inflation. So deducing real interest rates from nominal bonds requires at least the calculation of expected inflation, which is not trivial. The closest thing that we have in the data for real interest rates is inflation-adjusted government bonds. Such data is available for the UK and the US. In the UK,

[7]Yet, a utility with a time-varying degree of risk aversion may not be considered the most standard rational utility function.

inflation-adjusted government bonds (inflation-adjusted GILTs) have been available since the 1980s. In the US, the corresponding securities are called TIPS (Treasury Inflation-Protected Securities) and corresponding price data are available for roughly twenty years (Gürkaynak, Sack and Wright 2010).[8] A severe limitation of TIPS is that they are not as liquid as normal US treasuries. For this reason, this paper uses term premia measures produced by d' Amico *et al.* (2018) who computed risk-neutral yields and term premia, after taking account of the liquidity premia of TIPS over normal US treasuries.[9] As can be seen in Figure 2, in some periods liquidity premia of TIPS are considerable. Nevertheless, as shown in Figure 1, term premia are still significantly time-varying.

3.2 Real rates as a component of nominal rates

Figure 2 shows real yields at the top and nominal yields at the bottom. The plot reveals several key conclusions. Firstly, both nominal and real interest rates are time-varying. In addition, different maturities have different yields and the term structure seems to be upward-sloping in both cases. In other words, longer maturities are associated with higher yields. The slope of the term structure is also not constant, as the spread between yields of different maturities varies. Secondly, it is clear from Figure 2 that nominal rates are highly correlated with real rates. Considering the Fisher equation:

$$y_t^{nom,m} \approx y_t^{real,m} + E[\pi_{t,t+m}] \tag{1}$$

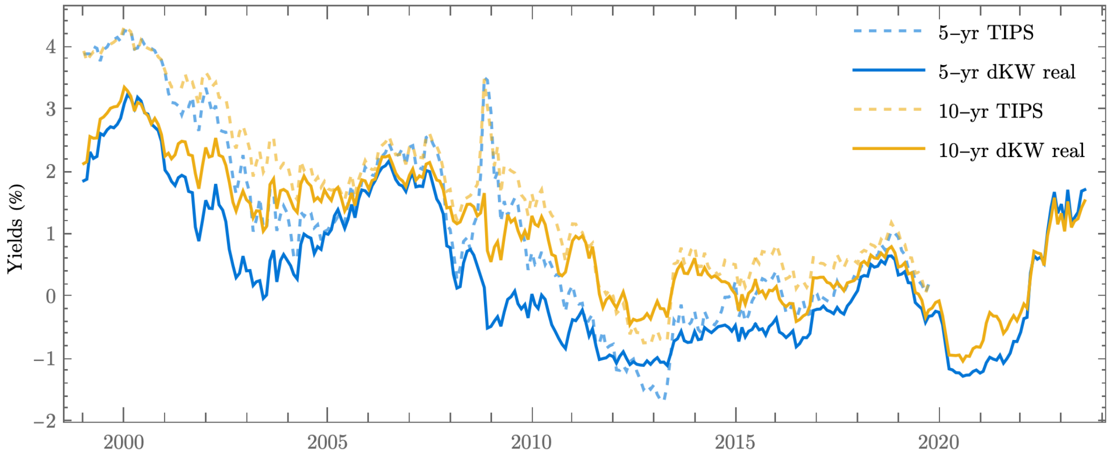
where m denotes the maturity of the underlying bond. The nominal rate ($y_t^{nom,m}$) can be thought of as a composite rate that includes two separate components, the real rate $y_t^{real,m}$, and expected inflation $E[\pi_{t,t+m}]$.[10] Figure 2 also shows that real interest rates are a significant and non-trivial component of nominal interest rates. Namely, real rates are moving substantially and mostly in parallel to nominal rates. Appendix B statistically verifies that the information contained in the movements

[8] Gürkaynak *et al.* (2010) has provided data starting from 1999. However, the full set of maturities is provided starting in 2002.

[9] Apart from liquidity issues related to TIPS, there is also a small concern (mostly with recently issued TIPS) that negative inflation is not correctly accounted for. This is because TIPS are guaranteed to pay investors at least the original principal value of the bond, even if the rate of inflation is negative. This makes inflation adjustment somewhat skewed. However, the effect will probably be small for securities that were issued several years prior, given that likely some inflation has already occurred and the probability that negative inflation will overcome it is small. Lastly, the accuracy of inflation adjustment can be debated, as the consumer price index might not capture the specific inflation concerns of investors.

[10] The Fisher equation can be made into an equality by adding an inflation risk premium.

of real rates explains a large proportion of the variation of nominal rates.



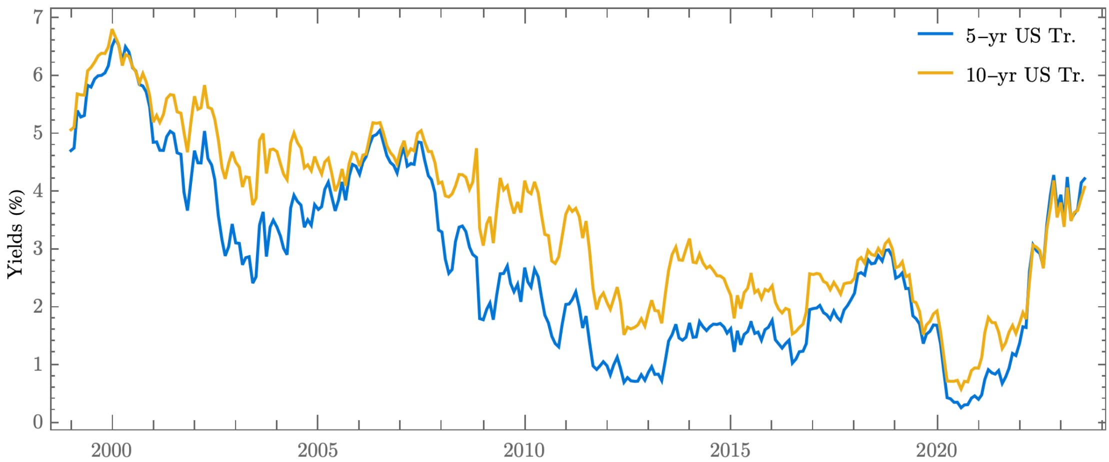


Figure 2: **Yields of US Treasuries**
TIPS data is taken from Gürkaynak *et al.* (2010), normal US treasury yields data is taken from Gürkaynak *et al.* (2007), and dKW real yields are taken from d' Amico *et al.* (2018). Real yields are the sum of risk-neutral yields and the real term premia. The difference between the dashed and solid lines corresponds to the liquidity premia of the TIPS over the normal treasuries. Normal treasuries are assumed as perfectly liquid.

3.3 Empirical evidence regarding term premia

Empirical research has predominantly focused on nominal bonds in relation to concepts such as term premia, *return predictability*, the *expectations hypothesis* (EH), and *excess volatility*. Specifically, predictability in nominal rates has been found by Fama and Bliss (1987) and Singleton (1980), who showed that yield spreads can partially predict excess returns of bonds over extended periods. This implies both the existence of term premia and that they are time-varying. In addition, this is equivalent to a violation of the EH, which has been verified by Cochrane and Piazzesi (2005) among others. The existence of excess volatility (Shiller 1979) also indicated time-varying term premia, because the excess volatility is evidence that changing economic conditions affect the value of long-term bonds beyond what can be explained by movements in expected short rates. Even though the literature has focused less on real term premia, relatively recent studies have concluded that real term premia are also positive and time-varying, after accounting for liquidity premia. In particular, Abrahams *et al.* (2016) estimated the five-to-ten-year real forward term premium and found that it has ranged roughly from 0% to 4% between 2000 and 2014 (Figure 14 in Appendix E).[11] d' Amico *et al.* (2018) found that the five-to-ten-year forward term premium has ranged roughly from -0.5% to 4% between 1980 and 2022 (Figure 1).[12] Pflueger and Viceira (2016) have demonstrated the existence of predictability of real excess returns, also implying the existence of time-varying real term premia. The conclusion that real term premia are substantial and time-varying is significant because it implies that models that exclusively use inflation processes are not sufficient to explain the dynamics of either nominal or real interest rates.

4 The Framework

A consumption-based framework in continuous time is used, which can accommodate a range of model variations. While consumption is used throughout the model variations, the framework could be reformulated in terms of wealth. So, wealth is considered as roughly interchangeable with consumption. In addition, financial intermediaries or general financial institutions do not technically consume or have wealth, but consumption could correspond to dividends, and wealth could correspond to total assets or market value. The framework is built upon three main components from which everything else is derived: 1) an exogenous consumption process, 2) a utility specification, and 3) a process for the state variable. The state variable determines the state of the economy, and it is either connected with some component of the consumption process or with some components of the

[11] Shown in Figure 5 of Abrahams *et al.* (2016).

[12] For the period between 2000 and 2014, the results of d' Amico *et al.* (2018) implied a somewhat smaller variability of term premia compared to the results of Abrahams *et al.* (2016).

utility function. Specifically, in the variations in this paper, the state variable is either connected to consumption trend (otherwise referred to as CD) or connected to consumption volatility (CV), or connected to the external habit of the utility function. These three options in combination with different calibrations and utility specifications give rise to a long list of variations and interpretations. To keep things simple, only one state variable is used for each model variation. Utility is either time-separable (TSU) or recursive (RU), in the latter case following Duffie and Epstein (1992).

4.1 Naming the variations

As mentioned already, several model variations are analysed in the main text of this paper, and many more in Appendix F. While models variations are explained in detail in Sections 4 and 5, for convenience, abbreviations for the model variations in Table 1 are also provided:

Model Variation Description	**Abbreviation**
Time-varying CD with time-separable utility.	TSU-CD
Time-varying CV with time-separable utility.	TSU-CV
Time-varying habit with time-separable utility.	TSU-Habit
Time-varying CD with recursive utility.	RU-CD
Time-varying CV with recursive utility.	RU-CV
Time-varying CD and CV with recursive utility.	RU-Mixed
High time-varying CV with positive correlation $\rho_{cx} > 0$, and time-separable utility.	TSU-HCV
Arbitrageur case with short-term rate decreasing in the investment opportunity and positive correlation $\rho_{cx} > 0$, with time-separable utility.	Arb-DP
Arbitrageur case with short-term rate decreasing in the investment opportunity and negative correlation $\rho_{cx} > 0$, with time-separable utility.	Arb-IN

Table 1: **Names of main model variations**.
The models are explained in Section 5.

4.2 Consumption process

Although consumption is often considered a fundamental choice variable for economic agents, it is assumed to be exogenous in this paper.[13] This approach is consistent with consumption having been decided at some earlier stage that is not explicitly modeled, and it significantly simplifies the analysis. In the most general form, consumption (C_t) follows the stochastic process expressed below:[14]

$$\mathrm{d}\log(C_t) = \mathrm{d}c_t = \mu_{ct}\mathrm{d}t + \sigma_{ct}\mathrm{d}W_{ct} \tag{2}$$

μ_{ct} denotes the CD at time t and σ_{ct} is the volatility coefficient of consumption growth at time t, which is multiplying the stochastic component $\mathrm{d}W_{ct}$.[15] In the remainder of the paper, CV refers to σ_{ct}.

4.3 Utility

Lifetime utility at time 0 takes the following form depending on the utility specification:

$$\underbrace{U_0 = \mathrm{E}_0 \int_0^{\infty} e^{-\rho t} u(C_t, S_t)\mathrm{d}t,}_{\text{TSU}} \qquad \underbrace{V_0 = \mathrm{E}_0 \int_0^{\infty} F(C_t, V_t)\mathrm{d}t}_{\text{RU}} \tag{3}$$

In both cases, there is an infinite horizon, with ρ representing the time preference parameter. In the case of TSU, flow utility u depends on the consumption flow (or just consumption for simplicity) and potentially on the surplus consumption ratio S_t, which is connected to the external habit.[16] In the variations without habit, S_t is taken to be equal to 1. In the case of RU, the aggregator function FF depends on consumption and on the current lifetime utility V_t which in the context of RU

[13] This is a standard choice in this literature. See for example Campbell and Cochrane (1999) and Bansal and Yaron (2004).

[14] It should be noted that for all model variations the same parameter symbols are used, and they should be distinguished by context. For example, in TSU-CD μ_{ct} is time-varying and a function of the state variable, while in TSU-CV μ_{ct} is a constant. The same applies to the symbols: σ_{ct} and σ_{xt}.

[15] W_{ct} is a standard Wiener Process associated with consumption such that $W_{ct} - W_{cs} \sim \text{Normal}(0, s-t)$.

[16] In the habit model of Campbell and Cochrane (1999), which is followed here, this variable is actually equal to $(C^a_t - X_t)/C^a_t$, where X_t is the level of habit and C^a is aggregate consumption.

is referred to as the value function. u and F take the following form:

$$\underbrace{u(C_t, S_t) = \frac{(C_t S_t)^{1-\gamma} - 1}{1-\gamma}}_{\text{TSU}}, \qquad \underbrace{F(C_t, V_t) = \frac{\rho\,(1-\gamma)\,V_t}{1-1/\psi}\left(\left(\frac{C_t}{((1-\gamma)\,V_t)^{-1/(1-\gamma)}}\right)^{1-1/\psi} - 1\right)}_{\text{RU}} \tag{4}$$

γ is the risk aversion parameter, and in the standard TSU case it is equal to relative risk aversion, which also equals the inverse of the IES. ψ is the IES parameter in the RU case.[17]

4.4 State variable process

At time t, the state variable x_t follows the process:

$$\mathrm{d}x_t = -\log(\phi)(\mu_{x0} - x_t)\mathrm{d}t + \sigma_{xt}\mathrm{d}W_{xt} \tag{5}$$

This expression describes an autoregressive stochastic process that reverts to the stochastic steady state (or steady state for simplicity), μ_{x0}.[18] The rate of reversion to the steady state is governed by ϕ, which is constrained to be between 0 and 1. Thus, $\log(\phi)$ is non-positive and it implies that when $x_t > \mu_{xt}$ ($x_t < \mu_{xt}$) the drift is downward-sloping (upward-sloping), always towards the steady state. $\mathrm{d}W_{xt}$ is also a standard Wiener process, and σ_{xt} is the volatility coefficient of the state variable and it is either a constant or it also depends on x_t. $\mathrm{d}W_{xt}$ can be correlated with $\mathrm{d}W_{ct}$, and the value of the correlation is captured by ρ_{cx}. In economic terms, the state variable plays a different role for each model variation. The full dependence of the model variations on the state variable is shown in Table 2: In some variations the steady state is at $x_t = 0$, while in others it is at $x_t = 1$, and x_t is positive with probability 1. This specification is used for the variations in which CV σ_{ct}, is proportional to the state variable, to ensure that σ_{ct} is positive.

4.5 Stochastic discount factor

4.5.1 Time-separable utility Case

In the TSU case, the stochastic discount factor (SDF) is the derivative of the flow utility with respect to consumption. In the general case the formula is the

[17] f has the form of a normalised aggregator as in Duffie and Epstein (1992).

[18] The steady state $x_t = \mu_{x0}$ does not necessarily coincide with the ergodic mean or median of the process when the diffusion of the process is not symmetric around the steady state value.

Model variation			
TSU-CD:	$\mu_{ct} = \mu_{c0} + x_t$	$\sigma_{xt} = \sigma_{x0}$	$\mu_{x0} = 0$
TSU-CV:	$\sigma_{ct} = \sigma_{c1} x_t$	$\sigma_{xt} = \sigma_{x1}\sqrt{x_t}$	$\mu_{x0} = 1$
TSU-Habit:*	$S_t = \bar{S} e^{x_t}$	$\sigma_{xt} = \sigma_{ct}\lambda(x_t)$	$\mu_{x0} = 0$
RU-CD:	$\mu_{ct} = \mu_{c0} + x_t$	$\sigma_{xt} = \sigma_{x0}$	$\mu_{x0} = 0$
RU-CV:	$\sigma_{ct} = \sigma_{c1}\sqrt{x_t}$	$\sigma_{xt} = \sigma_{x1}\sqrt{x_t}$	$\mu_{x0} = 1$
TSU-HCV:	$\sigma_{ct} = \sigma_{c1} x_t$	$\sigma_{xt} = \sigma_{x1}\sqrt{x_t}$	$\mu_{x0} = 1$
Arb-IN:	$\mu_{ct} = \mu_{c1} x_t^{1/4}, \sigma_{ct} = \sigma_{c1}\sqrt{x_t}$	$\sigma_{xt} = \sigma_{x1}\sqrt{x_t}$	$\mu_{x0} = 1$
Arb-DP:	$\mu_{ct} = \mu_{c1} x_t^{3/2}, \sigma_{ct} = \sigma_{c1} x_t^{2/3}$	$\sigma_{xt} = \sigma_{x1}\sqrt{x_t}$	$\mu_{x0} = 1$

Table 2: **Dependence on the state variable for each model variation**
* Following Campbell and Cochrane (1999) and Wachter (2006), the exact form of $\lambda(\cdot)$ is:

$$\sigma_{xt} = \sigma_{ct}\lambda(x_t) = \begin{cases} \sigma_{ct}\left(\frac{\sqrt{1-2x_t}}{\bar{S}} - 1\right) & \text{if } x_t < \frac{1-\bar{S}^2}{2} \\ 0 & \text{if } x \geq \frac{1-\bar{S}^2}{2} \end{cases}, \quad \bar{S} = \sqrt{\frac{\gamma}{-\log(\phi) - b/\gamma}} \tag{6}$$

following:

$$\Lambda_t = e^{-\rho t}(C_t S_t)^{-\gamma} \tag{7}$$

where S_t is only relevant in the habit model. Using the above expression, along with the consumption process (Equation 2) and the state variable process (Equation 5), Ito's Lemma can be implemented to get the stochastic differential equation (SDE) of the SDF:

$$\begin{aligned} \frac{\mathrm{d}\Lambda_t}{\Lambda_t} =& \Big(-\rho - \gamma\mu_{ct} + \frac{\gamma^2}{2}\sigma_{ct}^2 \Big)\mathrm{d}t - \gamma\sigma_{ct}\mathrm{d}W_{ct} \\ & \underbrace{+\big(-\gamma\log(\phi)x_t + 2\rho_{cx}\sigma_{ct}\sigma_{xt} + \sigma_{xt}^2 \big)\mathrm{d}t - \gamma\sigma_{xt}\mathrm{d}W_{xt}}_{\text{habit model only}} \end{aligned} \tag{8}$$

For the details of the derivation, see Appendix G.1.

4.5.2 Recursive utility case

In the case of RU, the stochastic process of the SDF is derived from the expressions for the value function and the aggregator function. The latter is given in Equation 4, and as shown by Tsai and Wachter (2018), the value function follows:[19]

$$V_t = \frac{C_t^{1-\gamma} e^{(1-\gamma)K(x_t)}}{1-\gamma} \tag{9}$$

[19] Similar results are common in the literature, see for example Benzoni, Collin-Dufresne and Goldstein 2011; Kraft, Seiferling and Seifried 2017.

V_t increases with K, which is a specific function of x_t that captures the full dependence of the value function on the state variable. At the end of this section, the expression above is justified, and a novel perturbation approximation that provides a formula for K is provided. Given the expression for the value function, Ito's Lemma can be implemented to get to the SDE of the SDF. The calculation here follows Chen, Cosimano, Himonas and Kelly (2009). In particular, the fundamental relationship is:

$$\frac{\mathrm{d}\Lambda_t}{\Lambda_t} = F_V(C_t, V_t)\mathrm{d}t + \frac{\mathrm{d}F_C(C_t, V_t)}{F_C(C_t, V_t)} \tag{10}$$

F_C and F_V denote partial derivatives of F with respect to consumption and the value function respectively. The first term on the right-hand side is the derivative of the flow utility with respect to the value function. The second term can be computed by applying Ito's lemma on the derivative of flow utility with respect to consumption.[20] The result is the following:

$$\begin{aligned}\frac{\mathrm{d}\Lambda_t}{\Lambda_t} =& \Bigg(\frac{\rho(-(1-\gamma\psi)e^{\frac{(1-\psi)K[x_t]}{\psi}} - \gamma\psi + \psi)}{1-\psi} - \gamma\mu_{ct} + \frac{\gamma^2\sigma_{ct}^2}{2} + \frac{\gamma(\gamma\psi-1)\rho_{cx}\sigma_{xt}\sigma_{ct}K'(x_t)}{\psi} \\ &+ \frac{(\gamma\psi-1)(2\psi(\mu_{x0}-x_t)\log(\phi)K'(x_t) + \sigma_{xt}^2((\gamma\psi-1)K'(x_t)^2 - \psi K''(x_t)))}{2\psi^2}\Bigg)\mathrm{d}t \\ &- \frac{(\gamma\psi-1)\sigma_{xt}K'(x_t)}{\psi}\mathrm{d}W_{xt} - \gamma\sigma_{ct}\mathrm{d}W_{ct}\end{aligned} \tag{11}$$

Now ψ can be replaced with $1/(1-\epsilon)$:

$$\frac{\mathrm{d}\Lambda_t}{\Lambda_t} = \left(\frac{\rho(-(1-\gamma\frac{1}{1-\epsilon})e^{\frac{(1-\frac{1}{1-\epsilon})K[x_t]}{\frac{1}{1-\epsilon}}} - \gamma\frac{1}{1-\epsilon} + \frac{1}{1-\epsilon})}{1-\frac{1}{1-\epsilon}} - \gamma\mu_{ct} + \frac{\gamma^2\sigma_{ct}^2}{2} + \frac{\gamma(\gamma\frac{1}{1-\epsilon}-1)\rho_{cx}\sigma_{xt}\sigma_{ct}K'(x_t)}{\frac{1}{1-\epsilon}}\right)\mathrm{d}t \tag{12}$$

The details of the derivation can be found in Appendix G.2. It is notable that in the special case of $\gamma = 1/\psi$, which corresponds to TSU, the equation above simplifies to the TSU formula in Equation 8. Also, the stochastic component relating to consumption ($-\gamma\sigma_{ct}\mathrm{d}W_{ct}$), is exactly the same as in TSU, and there is an extra component, namely $-\frac{(\gamma\psi-1)\sigma_{xt}K'(x_t)}{\psi}\mathrm{d}W_{xt}$, due to the direct dependence of the SDF on the state variable.

[20] This operation is performed by substituting the value function using Equation (9) and applying Ito's lemma based on consumption and the state variable as independent variables.

4.6 Instantaneous short-term rate

From the SDF the short-term rate is derived as follows:

$$\text{TSU: } r(x_t)\mathrm{d}t = -\mathrm{E}_t\left[\frac{\mathrm{d}\Lambda_t}{\Lambda_t}\right] = \Big(\rho + \gamma\mu_{ct} - \frac{\gamma^2}{2}\sigma_{ct}^2\Big)\mathrm{d}t \underbrace{+\big(\gamma\log(\phi)x_t - 2\rho_{cx}\sigma_{ct}\sigma_{xt} - \sigma_{xt}^2\big)\mathrm{d}t}_{\text{habit model only}}$$

$$\begin{aligned}\text{RU: } r(x_t)\mathrm{d}t = -\mathrm{E}_t\left[\frac{\mathrm{d}\Lambda_t}{\Lambda_t}\right] = \\ \frac{\rho\left((1-\gamma\psi)e^{\frac{(1-\psi)K[x_t]}{\psi}} + \gamma\psi - \psi\right)}{1-\psi} + \gamma\mu_{ct} - \frac{\gamma^2\sigma_{ct}^2}{2} - \frac{\gamma(\gamma\psi-1)\rho_{cx}\sigma_{xt}\sigma_{ct}K'(x_t)}{\psi} \\ - \frac{(\gamma\psi-1)\left(2\psi(\mu_{x0}-x_t)\log(\phi)K'(x_t) + \sigma_{xt}^2\left((\gamma\psi-1)K'(x_t)^2 - \psi K''(x_t)\right)\right)}{2\psi^2}\end{aligned} \tag{13}$$

In the standard TSU case, the short rate depends on three components. The first is the time preference parameter ρ. The second is $\gamma\mu_{ct}$, and it relates to the consumption smoothing motive. As CD increases, agents try to borrow to increase current consumption, and in equilibrium the short rate increases. The third is $-\gamma^2\sigma_{ct}^2/2$ and it relates to the precautionary savings motive. As consumption becomes more risky, agents try to save, and in equilibrium the short rate decreases. In TSU-Habit, there are extra components that relate both to the consumption smoothing motive and the precautionary savings motive, and they are due to the state variable being part of the utility function. Thus, as the surplus consumption ratio falls, marginal consumption increases even more than in standard TSU. So, the agent has an even higher motive to smooth consumption. However, in the same state of the world, the surplus consumption ratio is also much more volatile and the agent also has a higher precautionary savings motive. In Campbell and Cochrane (1999) these two opposite effects on the short-term rate are regulated by a parameter denoted b. If $b = 0$, then the short rate becomes a constant. If $b > 0$ ($b < 0$), then the short rate is decreasing (increasing) in the surplus consumption ratio.

In the RU case, the short rate becomes more complicated. However, for the main calibrations, the dominating additional effect comes from the fact that the marginal utility of consumption is expected to change as the state variable changes. As a consequence short rates are affected less by the consumption smoothing effect and the precautionary savings effect, and short rates under RU are less sensitive to the state variable than short rates under TSU.

4.7 Long-term bond

4.7.1 Bond pricing equation

Next, given the process of the SDF, the price of the long-term bond Q can be computed in the same way for both TSU and RU cases. The bond price is a function of the state variable x_t and its remaining maturity m. Thus, by using Ito's Lemma, the stochastic process follows:[21]

$$\mathrm{d}Q(x_t, m) = \big(-\log(\phi)(\mu_{x0} - x_t)Q_x - Q_m + \frac{1}{2}\sigma_{xt}^2 Q_{xx}\big)\mathrm{d}t + \sigma_{xt}Q_x \mathrm{d}W_{xt} \tag{14}$$

In the equation above, subscripts $\cdot_x$ and $\cdot_m$, denote partial derivatives with respect to the corresponding variable. The next step is to derive the partial differential equation (PDE) that Q obeys in these models. Thus, the pricing equation is used, following the approach in Cochrane (2009) and Chen, Cosimano and Himonas (2010):

$$\mathrm{E}\big[\mathrm{d}(\Lambda_t Q)\big] = 0 \Rightarrow \mathrm{E}\left[\frac{\mathrm{d}\Lambda_t}{\Lambda_t}Q + \mathrm{d}Q + \frac{\mathrm{d}\Lambda_t}{\Lambda_t}\mathrm{d}Q\right] = 0 \tag{15}$$

Substituting the expressions for Λ_t, $\mathrm{E}[\mathrm{d}\Lambda_t/\Lambda_t]$ and $\mathrm{d}Q$ from Equations (8), (13) and (14) respectively, gives rise to the PDE obeyed by Q:[22]

$$-Q_m - r(x_t)Q + \big(-\log(\phi)(\mu_{x0} - x_t) + A(x_t)\big)Q_x + \frac{\sigma_{xt}^2}{2}Q_{xx} = 0$$
$$\text{where: } A(x_t)\mathrm{d}t = \frac{\mathrm{d}\Lambda_t}{\Lambda_t}\mathrm{d}Q \tag{16}$$

The expression comprises five terms. The first is the derivative with respect to maturity Q_m. The second is the short rate term $r(x_t)Q$, which differs depending on the variation, as shown in Equation (13). The third is the expectation term $-\log(\phi)(\mu_{x0} - x_t)$, which captures the information that short rates may be expected to change in the future. The fourth is called the A term, and it is responsible for term premia, as it captures consumption-based risk.[23] The fifth is the diffusion term $\frac{\sigma_{xt}^2}{2}Q_{xx}$.[24] The solution of this equation is discussed next, while Appendix C shows in more detail how these five terms affect the term structure

[21] Given the flow utility function, investors' decisions are not affected by the level of consumption. This implies that the long-term bond is not going to be a function of consumption itself (see for example Tsai and Wachter 2018 who also commented on this).

[22] This equation is similar to a Black-Scholes equation.

[23] Risk is understood in the context of consumption-based asset pricing. Therefore, if the price of the bond does not co-vary with the SDF, then the A term is 0. The A term being 0 does not mean that the price of the bond is deterministic.

[24] This term is connected with the idea of *convexity* in finance.

of interest rates and their dynamics.

4.7.2 Solution of the pricing equation

Equation (16) is a PDE, and it is solved by use of the Feynman-Kac formula, which re-expresses the solution of a PDE as an expectation of a stochastic process. In particular, the solution of Equation (16) is:

$$Q(m, x_t) = \mathrm{E}_t\left[\exp\left\{\int_m^0 r(\tilde{x}_{t+s})\mathrm{d}s\right\}\right] = \mathrm{E}_t\left[\exp\left\{-\int_0^m r(\tilde{x}_{t+s})\mathrm{d}t\right\}\right] \tag{17}$$

where $\tilde{x}_0 = x_t$ and $\tilde{x}_t$ follow the modified stochastic process compared to the state variable:[25]

$$\mathrm{d}\tilde{x}_t = \Big(-\log(\phi)(\mu_{x0} - \tilde{x}_t) + A(\tilde{x}_t)\Big)\mathrm{d}t + \sigma_{xt}(\tilde{x}_t)\mathrm{d}W_{xt} \tag{18}$$

The expectation is computed using Monte Carlo simulations.

4.7.3 Risk-neutral yield and term premium

Instead of using a modified process, the original process of the state variable can also be used in the Feynman-Kac formula:

$$H(m, x_t) = \mathrm{E}_t\left[\exp\left\{\int_m^0 r(x_{t+s})ds\right\}\right] = \mathrm{E}_t\left[\exp\left\{-\int_0^m r(x_{t+s})\mathrm{d}t\right\}\right] \tag{19}$$

This is by definition the expected gross return from rolling over the short-term rate. Thus, $-\log(H(m, x_t))/m$ is by definition the risk-neutral yield, and the argument above shows that it corresponds to the solution of Equation (16), after setting $A(x_t) = 0$ for all x_t. In other words, the risk-neutral yield can be thought of as deriving from a bond priced by a risk-neutral investor. This also provides a natural way for computing term premia, which is:

$$TP(x_t, m) = \frac{-\log Q(x_t, m) - \big(-\log H(x_t, m)\big)}{m} \tag{20}$$

Namely, the term premium is the difference between the regular yield and the risk-neutral yield. Unfortunately, there is no analytic expression for term premia, given that Q and H are computed numerically. However, there is an analytic

[25] Here the dependence of σ_{xt} on $\tilde{x}_t$ is shown, in order to clarify that it is the same function as before, but it takes the modified variable as the argument. Using this modified process is equivalent to a change in the probability measure.

expression for function A in Equation 16, and it can serve as a diagnostic of term premia, as it is the component that distinguishes Q from H. Especially when the short-term rate is linear in the state variable, the sign of A determines the sign of term premia,[26] the time variability of A determines the time variability of term premia, and the size of A determines the size of term premia.[27] In addition, the size of A can easily be judged in comparison to the size of the expectation term $-\log(\phi)(\mu_{x0} - x_t)$, which also multiplies Q_x in the PDE. The expectation term and the A term are the two main drivers of the yield spread. Therefore, if the typical values of the expectation term are much larger than the typical values of the A term, this implies that the yield spread is due to expected changes in short-term rates in the future. On the other hand, if the values of the two terms are comparable in size, then the yield spread likely contains a component due to the term premium as large as a component due to the expectation term. This comparison is illustrated in practice in Section 5.

4.8 Perturbation approximation for K function in the recursive utility case

As mentioned in Subsection 4.5.2, given the process of the SDF it is possible to compute the price of bonds in the RU case in the same way as in the TSU case. This in turn requires an expression for the value function. This subsection is dedicated to explaining a novel perturbation method to approximate function K that was used in the RU value function. This novel method is explored in detail in Melissinos (2023).

Equation (7) shows how the value function can be written in a form that separates the dependence on consumption and the dependence on the state variable. Function K captures the dependence on the state variable, and it is the solution of the following ordinary differential equation (ODE):

$$-\frac{1}{2}\gamma\sigma_{ct}^2+\frac{1}{2}\sigma_{xt}^2\left(K''(x_t)-(\gamma-1)K'(x_t)^2\right)-\log(\phi)(\mu_{x0}-x_t)K'(x_t)+\frac{\rho\left(e^{-\epsilon K(x_t)}-1\right)}{\epsilon}+\mu_{ct}=0 \tag{21}$$

[26] In particular, term premia have the sign of the product of A with the derivative of Q with respect to the state variable x_t, which usually has the same sign as the derivative of the short rate with respect to the state variable x_t.

[27] To be precise, term premia depend on the entire pricing Equation (16). However, if the short-term rate is linear and the effect of the diffusion term is small, then the bulk of the time-varying behaviour of term premia is determined by A. In the explanation provided here, the diffusion term and the non-linearities are assumed to have a small effect on the yield spread. A detailed analysis is conducted in Appendix C.

where the substitution $\psi = \frac{1}{1-\epsilon}$ has been made. For $\psi = 1$ ($\epsilon = 0$) Equation (21) has an analytic solution. This can then be used to create a global perturbation solution in terms of the state variable, which is expressed in terms of ϵ.[28] In particular, Equation (21) can be expanded to:

$$-\frac{1}{2}\gamma\sigma_{ct}^2 + \frac{1}{2}\sigma_{xt}^2(-(\gamma-1)\left(\epsilon^2 K_2'(x_t) + \epsilon K_1'(x_t) + K_0'(x_t)\right)^2 + \epsilon^2 K_2''(x_t) + \epsilon K_1''(x_t) + K_0''(x_t)) + \mu_{ct}$$
$$+\frac{\rho\left(e^{-\epsilon\left(\epsilon^2 K_2(x_t)+\epsilon K_1(x_t)+K_0(x_t)\right)} - 1\right)}{\epsilon} - \log(\phi)(\mu_{x0} - x_t)\left(\epsilon^2 K_2'(x_t) + \epsilon K_1'(x_t) + K_0'(x_t)\right) \approx 0 \tag{22}$$

Here function K has been expanded up to second order, but it could also be expanded further. Given this expansion, the equation admits a solution of the form:

$$\begin{aligned}
K_0(x_t) &= a_{0,0} + a_{0,1}x_t \\
K_1(x_t) &= a_{1,0} + a_{1,1}x_t + a_{1,2}x_t^2 \\
K_2(x_t) &= a_{2,0} + a_{2,1}x_t + a_{2,2}x_t^2 + a_{2,3}x_t^3 \\
&\cdots
\end{aligned} \tag{23}$$

This solution can be plugged into the ODE (22), and for each m, n, $a_{m,n}$ can be derived by setting each factor of $x_t^m \epsilon^n$ equal to zero. This leads to a linear equation for each coefficient.[29] Conveniently, these equations can be solved successively so that for each equation there is only one unknown. As can be seen from equation (23), the full solution is a sum of polynomials in terms of x_t. For each successive order of ϵ, the order of the polynomial increases by one. While it is possible to compute many orders of approximation, eventually the computation becomes expensive, as each order of ϵ requires the solution of more linear equations, and each equation has an increasing complexity.

The solution in Tsai and Wachter (2018) only derived $K_0(\cdot)$ which is the first term in formula (23) and it is the "zeroth" order approximation in terms of ϵ or equivalently ψ. My approximation is useful because it allows a much larger range of values for ψ, and it provides an analytic expression that is easy to include

[28]The approximation is global in terms of the state variable x_t, as the perturbation is done with respect to parameter ϵ. Nevertheless, it is not valid for all values of x_t. In particular, the approximation takes such a form, so that its validity depends on different regions of the state variable. In the region where it converges, the quality of the approximation is high for all values of x_t, but outside this region, the series diverges.

[29]Apart from coefficient $a_{0,1}$ which may require the solution of a second order equation.

in the Monte Carlo simulations, that solve the pricing equation. Given that the solution provided by Tsai and Wachter (2018) is analytically correct only for $\epsilon = 0$ ($\psi = 1$), implementing the method for other values of ψ is not easily justifiable, even if in practice it would generate qualitatively or even quantitatively similar results. It should also be noted that the full perturbation series corresponds to the exact solution to the ODE, even if the series diverges. Namely, it is the unique perturbation series that represents the solution. Thus, it is highly likely that with some extra mathematical analysis, it can be re-expressed in terms of known special functions, and we can get an exact answer that is practically trivial to compute for arbitrary order. Nevertheless, the method in its current form allows the researcher to easily approximate the value function, while also practically checking its convergence. The value function can then be used in the pricing equation to directly get the price of assets while being robust to a large range of parameters for the IES. Thus, this method can be implemented widely in RU models. My method is described in detail in Melissinos (2023).[30]

4.9 Calibration

General parameters	
Relative risk aversion γ	2.0
Rate of time preference ρ	0.012/yr
Steady state CD μ_{c0}	0.0252 /yr
Steady state CV σ_{c0}	0.02/yr
Steady state reversion $\log\phi$	log(0.92)/yr

Table 3: **Calibration of common parameters**

Given that the goal of the paper is to identify mechanisms that are consistent with the patterns of term premia in the data, the emphasis is not on providing a perfect calibration. Instead, the focus is on finding the combination of the utility specification and the consumption process that generates the observed patterns in term premia. Thus, several parameter choices are explored, and the calibration of each model variation is reported in the corresponding figure showing the results. Nevertheless, there are common parameter choices across model variations. These

[30] One limitation of the method, that this paper's contribution does not overcome, is that the parameters of the processes should be at most linear functions of the state variable. In particular, σ_{xt}^2 and σ_{ct}^2 are linear in the state variable. Therefore, unlike in the TSU case where $\sigma_{ct} \propto x_t$, here $\sigma_{ct} \propto \sqrt{x_t}$. This is investigated in more detail in Melissinos (2023), but the main implication is that CV is relatively restricted in its variability.

are reported in Table 3 and they follow Wachter (2013), while the risk aversion parameter γ is set equal to 2 following Wachter and Zhu (2019).

5 Results

5.1 How variations are evaluated

In evaluating the features of term premia, the model variations need to generate empirically plausible short rate volatility.[31] Higher short rate volatility can give rise to higher-term premia. Thus, a relatively large short rate volatility is assumed to give these variations the best chance of success. Their performance is compared, by plotting model-implied term premia as a function of the state variable next to the time series of estimated term premia (these were already shown in Figure 1). In making the comparison, the focus is more on the variability than the level.[32] If the models generate roughly the same pattern of variability, then they are considered a success. In most cases, it is obvious when the models fail to generate the patterns of term premia in the data.

Table 4 shows information for function A for six separate variations with moderate CV, and Figures 3 and 5 show the corresponding five and ten-year term premia, which are discussed in Subsections 5.2.1 and 5.2.2 respectively. Further variations and calibrations are shown in Appendix F. With these results, apart from illustrating the standard mechanisms, a helpful reference is also provided, as in the literature state-dependent term premia are rarely provided. Furthermore, Table 5 and Figure 6 show variations with high CV, whose consequences have not been analysed before with respect to term premia. This paper is the first to show that these variations can generate the features of term premia in the data. These latter results are discussed in Subsection 5.3. While for the main results estimated premia from d' Amico *et al.* (2018) are used, Abrahams *et al.* (2016) also estimated the term premium, and they provide a decomposition of the five-to-ten-year forward. Estimations from both papers are shown in Appendix E,[33] and

[31] The calibration of the state variable is described in Appendix D.

[32] For example, in the data, especially recently, term premia seem to also become negative. However, the successful models all have exclusively positive term premia. This is not a large issue, as the focus is on the variability of term premia and the models investigated here only have one state variable. In a full explanation of term premia and interest rates more generally, at least two variables would be necessary, given that a principal component analysis of the yields and spreads requires at least two principal components to explain the bulk of the variation (this is shown in Appendix B)

[33] The two measures are similar, with the term premium in Abrahams *et al.* (2016) reaching relatively higher values. In addition, the risk-neutral yield has much less variability in the

the five-to-ten-year forward term premium generated by the variations analysed in the current paper is shown in Appendix F.

5.2 Moderate consumption volatility

5.2.1 Time-separable utility

Model variation	$A(x_t)$	ρ_{cx}	Range of A term*	Range of expectation term*
TSU-CD	$-\gamma\rho_{cx}\sigma_c\sigma_x$	$+$	(-0.0002,-0.0002)	$(0.0023, -0.0022)$
TSU-CV	$-\gamma\rho_{cx}\sigma_{ct}\sigma_{xt}$	$-$	(0.0012, 0.0086)	(0.048, -0.052)
TSU-Habit	$-\gamma\rho_{cx}\sigma_c\sigma_{xt} - \gamma\sigma_{xt}^2$	$+$	(-0.083, -0.011)	(0.034, -0.030)
RU-CD	$-\gamma\rho_{cx}\sigma_c\sigma_x - \frac{(\gamma\psi-1)\sigma_x^2 K'(x_t)}{\psi}$	$+$	(-0.001250, -0.001253)	(0.0023, -0.0022)
RU-CV	$-\gamma\rho_{cx}\sigma_{ct}\sigma_{xt} - \frac{(\gamma\psi-1)\sigma_{xt}^2 K'(x_t)}{\psi}$	$-$	(0.0018, 0.0040)	(0.030, -0.034)
RU-Mixed	$-\gamma\rho_{cx}\sigma_{ct}\sigma_{xt} - \frac{(\gamma\psi-1)\sigma_{xt}^2 K'(x_t)}{\psi}$	$+$	(-0.0045, -0.010)	(0.033, -0.031)

Table 4:
Information on function A from Equation (16) in different model variations with moderate CV. The t-subscript has been dropped from the quantities that are not time-varying according to the variation.
* This range covers the typical values of the state variable. The values correspond to the dashed vertical lines in Figures 3 and 5.

As mentioned earlier, the three main mechanisms analysed are time-varying CD, CV, and surplus consumption ratio (in TSU-Habit).[34] The effect of these mechanisms on term premia can first be understood by looking at function A for each of the variations. Table 4 shows the functional form of A, and which components are time-varying. It also shows the typical range for the size of the A term and the expectation term. As mentioned earlier, the sign of A, in conjunction with the slope of the short-term rate, determines the sign of term premia, while the size and variability of A also determine the size and variability of term premia. In TSU-CD the short rate is increasing with CD, due to the consumption smoothing

Abrahams *et al.* (2016) estimation.

[34]The first two mechanisms can also be found in the long-run risk models introduced by Bansal and Yaron (2004), who used a RU.

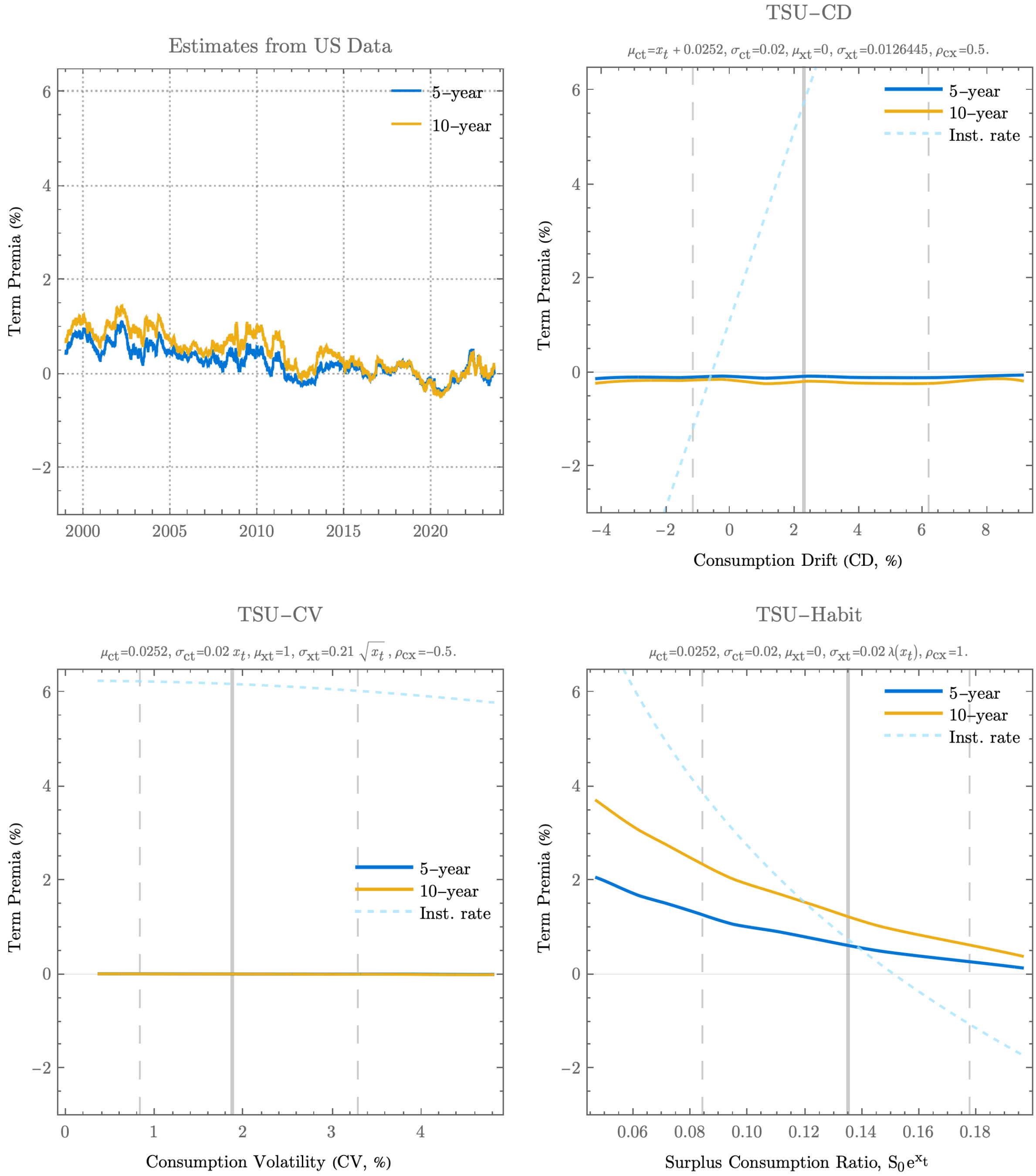


Figure 3: **Term premia in standard models with TSU**
The top left plot shows estimates of term premia according to d' Amico *et al.* (2018). The remaining plots show state-dependent term premia for three standard variations, namely variations with a) time-varying CD, b) time-varying CV, and c) an external habit in the utility function respectively. The dashed line shows the short-term rate.
The vertical dashed lines correspond to the typical values of the state variable based on simulations. The full range of the x_t-axis includes extreme values of the state variable, which are still possible (see Appendix D or Figure 16 for the exact definition of the ranges).

motive.[35] As a result, in conjunction with $\rho_{cx} > 0$ the term premia are negative and constant in the state variable. The intuition for negative term premia is that the short-term rate goes up and bond prices go down when CD rises, which is also the time that consumption tends to increase (due to $\rho_{cx} > 0$). This means that long-term bonds act as a hedge, and they command a negative term premium. Apart from the negativity of term premia, A typically takes much smaller values in absolute value compared to the expectation term, implying that term premia should be very small. Thus, instead of positive, time-varying and sizeable, as in the data, term premia are negative, constant and small. Alternatively, for TSU-CV, the short rate is decreasing in CV, due to the precautionary savings motive, and $\rho_{cx} < 0$. Therefore, the A term is positive and time-varying in the state variable, as it includes CV σ_{ct} (in this specification σ_{xt} is also time-varying). As a result, the term premia are again negative (they have the same sign as the slope of the short-term rate), but in this case, they are time-varying. However, the A term is much smaller in absolute value compared to the expectation term, so term premia apart from negative are again very small. Figure 3 shows the term premia for these two variations in comparison to the time series of term premia in the data.[36] It is evident from the figure that as the state of the economy changes, term premia would hardly move away from 0, and they would not be able to generate the variation estimated in the time-series. From the functional form of A it also follows that assuming a different sign for ρ_{cx}, would imply term premia of the opposite sign in both cases. However, for a representative consumption process it is reasonable that an increase of CD is associated with an increase in consumption itself, while an increase in CV is associated with a decrease in consumption.[37] In Appendix F the results above are verified for several different calibrations.

The mechanisms discussed above use the power utility setup. Here, the effect of including external habit in the utility function as in Campbell and Cochrane (1999)

[35]This means that the stochastic component of consumption is positively correlated with the stochastic component of the state variable, which is associated with CD. To avoid this long description, ρ_{cx} is mostly used.

[36]In TSU-CV the short rate is also insensitive to CV.

[37]This is intuitive if the consumption process is thought of as a relatively independent consumption process that determines the short-term rate. However, if the short-term rate is the independent variable, and the consumption process is reacting, then it makes sense that as the short-term rate decreases, borrowing becomes cheaper and consumption temporarily increases. This can can either imply that CD decreases, as consumption comes back to its normal level, or that CV increases as the agent has less savings. In both cases, the sign of ρ_{cx} is the opposite compared to the first scenario. It is conjectured that this should not happen in a large economy with a short-term rate determined by the behaviour of a representative agent. However, it could also be argued that the short-term rate is the independent force in the economy, due to the actions of the monetary authority.

is discussed. As shown by Wachter (2006), TSU-Habit can generate the basic patterns of term premia that we see in the data. As mentioned previously, models with time-varying risk aversion, are the only alternative to models with stochastic volatility for explaining real term premia. Thus, this model is analysed within the current setup, in order to comprehensively describe the possible explanations to real term premia, and delineate its main differences compared to the alternative explanation that is introduced in the next subsection. Table 4 shows that the habit model has an extra term in the functional form of A. It turns out that this second term is dominant because the state variable volatility is in most states much larger compared to CV ($\sigma_{xt} \gg \sigma_{ct}$).[38] As a result, the sign of A does not depend on ρ_{cx} (which in the canonical habit model is equal to 1 anyway, as consumption completely determines the habit variable.), and the sign of term premia is determined exclusively by the slope of the short-term rate as a function of the surplus consumption ratio. As discussed in Subsection 4.6, this relationship in TSU-Habit depends on parameter b, which is chosen positive so that the short-term rate is decreasing and the term premia are positive.[39] Furthermore, term premia are large, as the value of A is large compared to the expectation term. Lastly, term premia are time-varying, given that A includes σ_{xt}^2, which is time-varying. Namely, the variability of term premia is due to the heteroskedasticity of the state variable, which is amplified because A includes the square of the volatility.[40] Therefore, term premia are positive, time-varying, and large. This is explicitly shown in Figure 3, and the typical amount of variability, captured between the dashed lines, matches closely the variability in the estimated term premia.

5.2.2 Recursive utility

In this subsection, the results are extended to RU. This case is arguably of higher interest, as it separates risk aversion and IES. Moreover, Bansal and Yaron (2004) were able to use this feature in conjunction with time-varying CD and CV in long-run risk models, to explain the equity premium puzzle. Indeed, similar to TSU-Habit, as is shown in Table 4, the RU variations have an extra term in

[38]The size of the two terms is shown in the right plot of Figure 4

[39]This was also the choice of Wachter (2006), while Campbell and Cochrane (1999) set $b = 0$ in the final version of their paper (in an earlier version they also investigated $b > 0$). Appendix F also derives the results of a variation in which $b < 0$. In this case, the short rate is increasing in the surplus consumption ratio, and term premia are negative, time-varying and large in absolute value.

[40]Appendix F imposes homoskedasticity, and this leads to constant term premia. Admittedly, this is contrary to the spirit of the habit model.

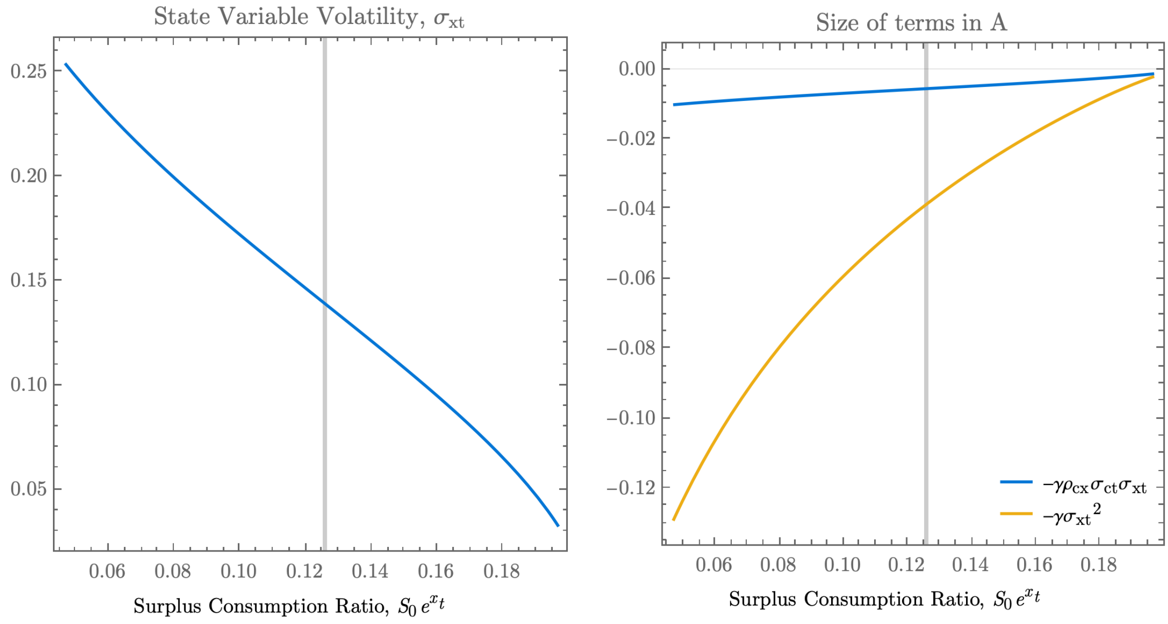


Figure 4: **Terms related to TSU-Habit**
The left plot shows the value of the volatility coefficient of the state variable in TSU-Habit. The right plot shows the magnitude of the two terms in the A function in TSU-Habit.

function A. This term disappears for $\gamma = 1/\psi$ which coincides with the special case in which utility becomes time-separable. Similar to TSU-Habit, this term dominates the A function. Thus, the sign of term premia does not depend on the sign of ρ_{cx}, but on the slope of K, which turns out to match the slope of the short-term rate both in RU-CD and in RU-CV. This means that negative term premia are now a more robust prediction compared to TSU-CD and TSU-CV. However, in the case of RU-CD A is significantly larger compared to TSU-CD. Therefore, term premia are negative, and constant, but can be somewhat sizeable in absolute value. In contrast, RU-CV shows the same patterns as TSU-CV. Given that the short rate hardly exhibits variability in RU-CV, RU-Mixed is also computed. This variation includes both time-varying CD and CV, governed by the same state variable. However, A in this variation is also quite small, and the term premia are small in absolute value. The term premia for RU-CD, RU-CV and RU-Mixed are shown in Figure 5, and it is clear that they cannot generate the variability in the estimated term premia. Appendix F has further variations with different calibrations verifying these results. Intuitively, RU models might be considered good candidates for explaining term premia due to their flexibility

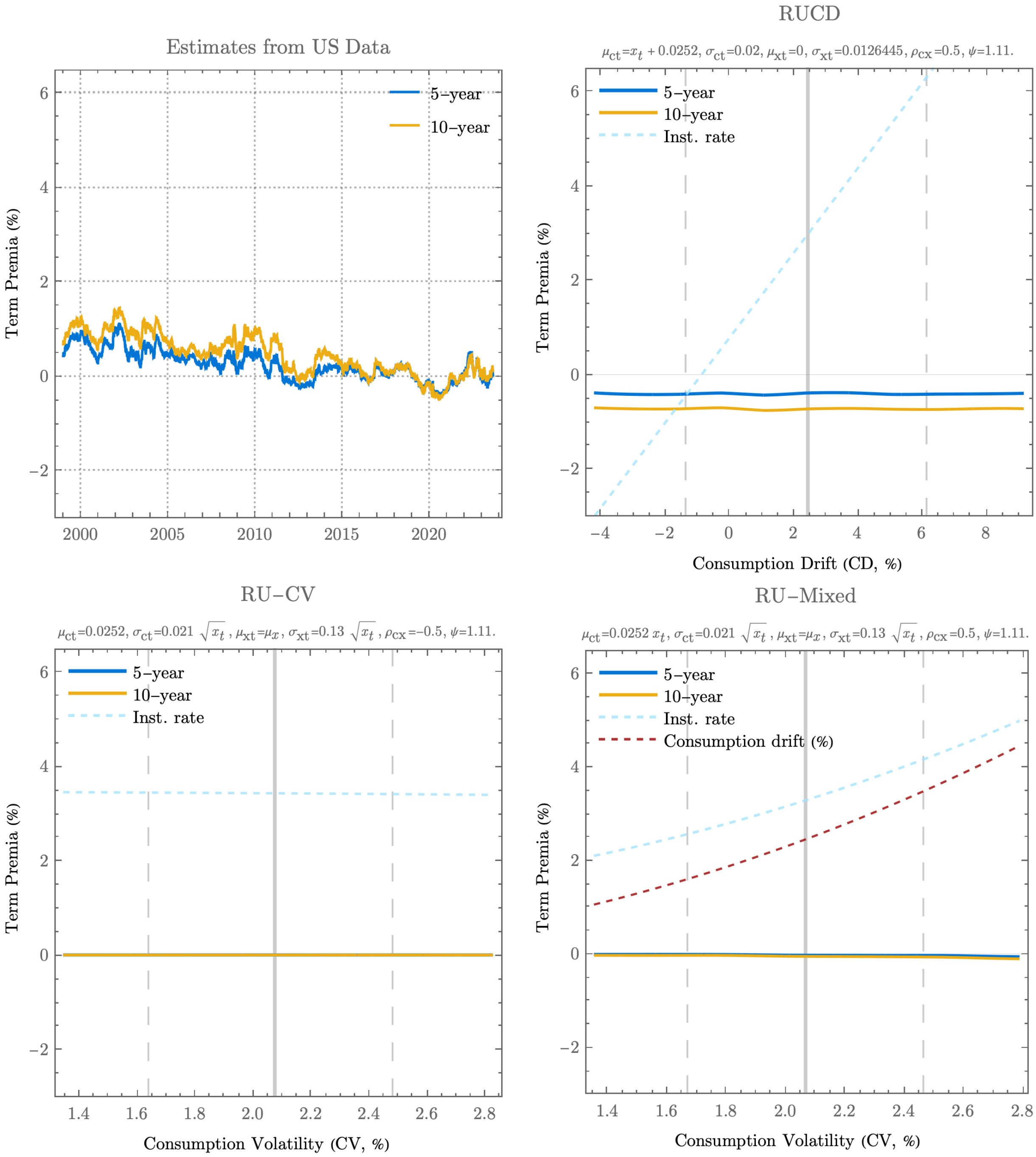


Figure 5: **Term premia in standard models with RU**
See Figure 3 for details. The variations are shown in the titles of the subplots.

in separating risk aversion and IES. However, term premia are constant in RU-CD and very small in RU-CV, while they are in both cases negative. This result is consistent with the literature. Specifically, Bansal and Shaliastovich (2013) study term premia in RU models, but they investigate the variability in *nominal* term premia and their mechanism involves inflation. Gomez-Cram and Yaron (2021) provided a similar explanation for nominal term premia using RU that also relies on inflation. Hence, the real term premia that they generate are not substantially time-varying. Van Binsbergen, Fernández-Villaverde, Koijen and Rubio-Ramírez (2012) also considered a RU setup with inflation, and they find that nominal term premia can be positive, for very high risk aversion values. However, they also found that real term premia are negative.

5.3 High consumption volatility

Model variation	$A(x_t)$	ρ_{cx}	Range of A term*	Range of steady state reversion term*
High CV	$-\gamma\rho_{cx}\sigma_{ct}\sigma_{xt}$	$+$	(-0.0079,-0.0057)	(0.0047, $-$0.0053)
Arb-IP	$-\gamma\rho_{cx}\sigma_{ct}\sigma_{xt}$	$-$	(-0.0069, -0.052)	(0.048, -0.052)
Arb-DN	$-\gamma\rho_{cx}\sigma_{c}\sigma_{xt}$	$+$	(0.0096, 0.045)	(0.047, -0.054)

Table 5:
Information on function A from Equation (16) in different model variations with HCV.
* This range covers the typical values of the state variable. The values correspond to the dashed vertical lines in Figure 6.

In general, agents should be independently adjusting their investment and consumption. Thus, given the same asset-pricing processes, if optimising agents are heterogeneous in their utility functions, they will have different consumption processes. Given a utility function, the consistency of term premia with the consumption process can be checked independently for each consumer. Previously, it was shown that representative consumer explanations of term premia require time-varying risk aversion. This raises the question of whether there is *any* consumer group whose consumption process is consistent with term premia, without assuming time-varying risk aversion. Given the negativity and the small size of term premia found in the previous subsection, it is reasonable to assume that the answer is again no. However, it turns out that other explanations rely on the dynamics of

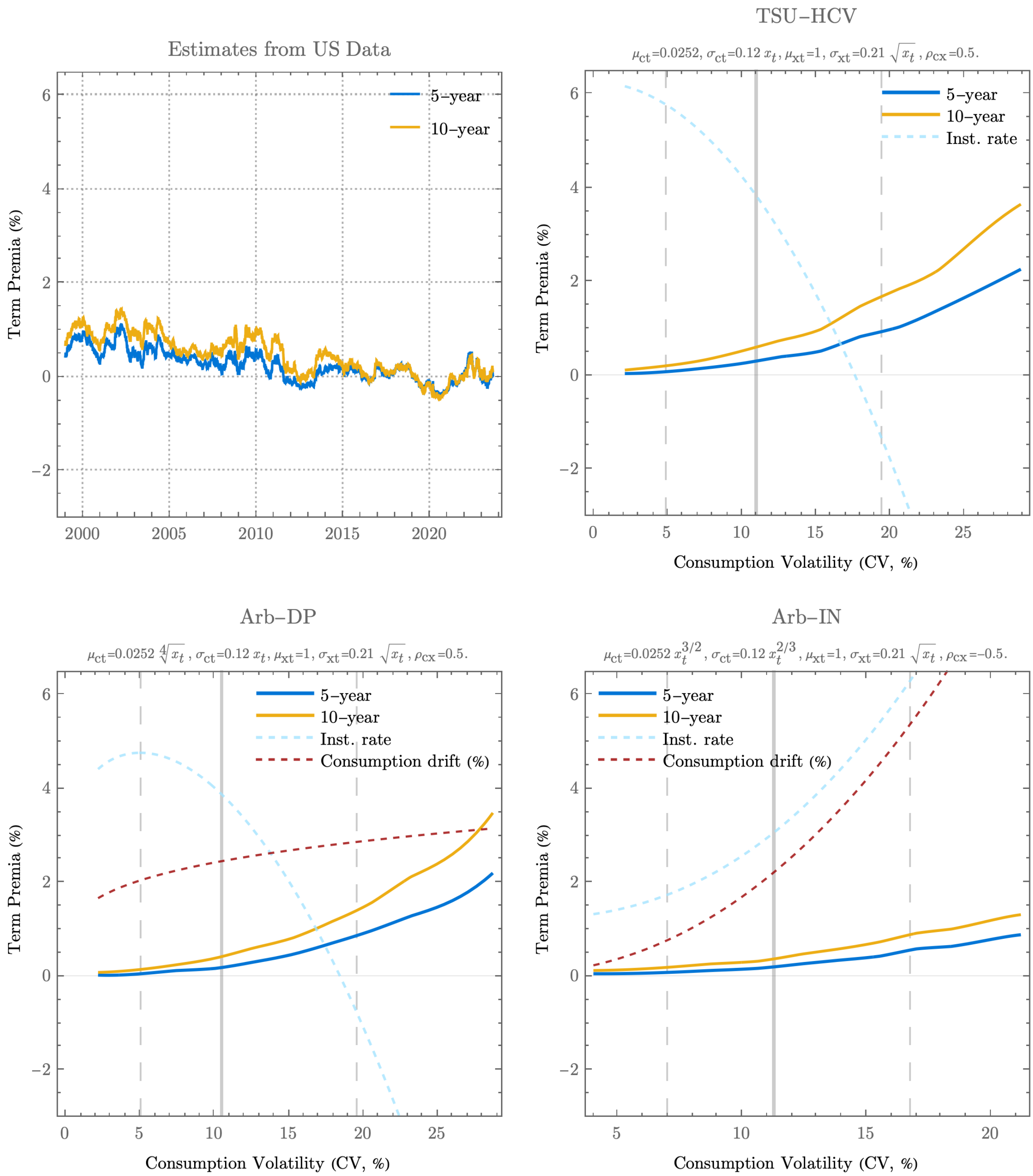


Figure 6: **Term premia in models with TSU and HCV**
See Figure 3 for details. The variations are shown in the titles of the subplots.

the consumption process within TSU. Table 3 shows information on function A for these cases, while Figure 6 shows the corresponding state-dependent term premia. As has been shown previously, time-varying CV implies time-varying term premia. Thus, starting from time-varying CV, the way forward is in principle simple based on the expression for A. By changing the sign of ρ_{cx} and increasing the steady state level of σ_{ct}, term premia become positive and large. Indeed, this works in generating the amount of variability in the estimated term premia (Figure 6). This is noteworthy given the difficulty encountered previously in generating any amount of significant time-varying term premia. However, are these two necessary modifications economically sensible?

As shown in Figure 6 the typical variability of the state variable (area between dashed lines) ranges from 5% to more than 20% CV per year. Even accounting for potential mismeasurement of aggregate consumption, this range is excessively large compared to the volatility observed in aggregate consumption data.[41] Rather the levels are similar to the levels of volatility in the stock market.[42] Therefore, this approach is not consistent with a representative consumer whose consumption coincides with aggregate consumption. However, this does not mean that the consumption process is too extreme for *any* consumer. Firstly, if financial markets are incomplete, and risk sharing is not possible, then idiosyncratic CV is relevant for asset prices (Constantinides and Duffie 1996). This means that aggregate CV could already be underestimating individual CV. Next, while 12% steady state CV is large compared to aggregate CV, it is not large compared to asset price volatility in financial markets. For people whose wealth lies in the financial sector, 12% wealth volatility is entirely plausible, and according to standard consumption-based portfolio theory, CV should follow wealth volatility. Lastly, there is also direct evidence that CV is much higher for some groups of consumers. While the paper does not take a position on whether the marginal investors are rich or poor, Ait-Sahalia, Parker and Yogo (2004) showed that the CV of rich individuals could be much higher compared to aggregate CV. In particular, while they reported that the annual standard deviation of non-durables and services was 2.3% according to the standard NIPA data, they measured an annual standard deviation of 19.6% for luxury retail sales and 20.4% for charitable contributions of wealthy individuals.[43] These values are both significantly larger than the steady state CV

[41]Savov (2011) suggested that due to mismeasurement, consumption volatility is underestimated.

[42]For example over the last nine years the standard deviation of annual returns of the SP500 was about 17.4%.

[43]NIPA refers to the national income and product accounts produced by the Bureau of Eco-

of the model variations in this subsection, which is equal to 12%.[44] Based on these results, Ait-Sahalia *et al.* (2004) also argued that the equity premium puzzle is less of a puzzle when considering the consumption process of rich consumers, as HCV also implies a sizeable equity premium.[45]. Similarly, Malloy, Moskowitz and Vissing-Jørgensen (2009) provided evidence that wealthy stockholders' CV is roughly three times higher compared to non-stockholders, while also showing that bond returns can be predicted by the covariance of wealthy stockholders' consumption growth with returns. Lastly, in a different strand of evidence, Carpenter, Demiralp, Ihrig and Klee (2015) also showed that during the conduct of unconventional monetary policy, it was households that traded with the Fed, when it was trying to affect long-term yields.[46] Beyond genuine consumers, as has already been mentioned, marginal investors could correspond to financial intermediaries or institutions whose market value could quite plausibly vary with a standard deviation that fluctuates between 5% and 20%. All this is consistent with the idea that a small group of investors with HCV are driving term premia.

The second required assumption for the mechanism is that ρ_{cx} is positive. Previously, it was argued that this is not plausible for a representative consumer, because an increase in consumption risk should induce consumers to consume less and save more. However, the consumer-investors in TSU-HCV could be a small part of the overall population, and in this case, $\rho_{cx} > 0$ can be justified. As CV increases, the short-term rate goes down, and this leads to an increase in bond prices. Thus, bondholders would then increase their consumption, given that their wealth also increases. An alternative intuition is that, as the short-term rate decreases, consumption increases due to borrowing, which in turn increases CV.

The effect of HCV on term premia has been shown only for $\gamma = 2$. Apart from further variations in Appendix F, Figure 7 shows the different levels of term premia on the same scale for various values of γ and various values of steady-state CV. The results are interesting in several ways. Firstly, it stands out that different values of γ lead to huge changes in term premia, when CV is high. This means that term premia in TSU-HCV are highly sensitive to risk aversion levels. On the other hand, term premia are so small when CV is low, that moderate increases in risk aversion are not able to generate the required variability. Thus, even if $\gamma = 4$,

nomic Analysis of the US Department of Commerce. Ait-Sahalia *et al.* (2004) also included other measurements on the sales of luxury retail products.

[44] The standard deviation of consumption growth calculated from simulations also takes values similar to the CV of the model.

[45] This is also shown in an example in Melissinos (2024).

[46] This household classification is likely somewhat different compared to what the label implies, for instance, it usually includes hedge funds.

CV needs to be able to reach at least 10%, so that time-variability in term premia is generated. This subsection shows how some consumers could have consumption processes that are consistent with the main features of term premia. By restricting attention to these investors, and not introducing a full heterogeneous agent model, many different variations can be examined. Nevertheless, it is important to also consider the potential behaviours of the remaining agents in the economy. For instance, they could be investing in the bond market, but their behaviour could be explained by more complicated or alternative models. It could also be the case that other investors in the bond market are entities, such as hedge funds and pension funds that are not appropriately modelled as consumers. The only requirement for the remaining investors is that they do not trade in such a way, that induces extensive risk sharing with HCV investors. If they did, then this would lead to a decrease in the CV of the HCV investors. Alternatively, many consumers may not be participating in the bond market at all.[47] In both cases the other agents can have moderate consumption processes, and be primarily responsible for aggregate consumption dynamics.[48]

TSU-HCV has been the simplest consumption-based variation that can generate large term premia. However, given a high CV, slightly more complicated variations can be examined, in which CD and CV are simultaneously changing. These are referred to as "arbitrageur" variations as in Vayanos and Vila (2021), who suggested that the term structure of interest rates is driven by "arbitrageurs", who take advantage of investment opportunities in the bond market. As these opportunities can be risky, arbitrageurs are not able to fully equate rates and eliminate the effect of the demand of idiosyncratic investors or "preferred habitat investors", as they are called in the article.[49] Here, the latter investors are not considered and the focus is placed on arbitrageurs. They are marginal investors in the bond market. Their consumption process should be consistent with the observed term structure of interest rates, including term premia. Thus, the consumption process of the arbitrageurs is assumed to have two main features. Firstly, their CV is high (similar to TSU-HCV). Secondly, as the investment opportunity increases, both CV and CD rise. This occurs because the higher investment opportunity offers higher expected returns, which implies a higher CD. At the same time, the higher

[47]Or they may not be marginal investors due to short-selling constraints. For instance, an investor who is constrained from shorting one end of the term structure could be holding some long-term bonds, but this does not make her a marginal investor of bonds in general.

[48]A fuller analysis would provide a full heterogeneous agent model explaining to what extent idiosyncratic consumption risk can be insured through financial markets.

[49]Given that there is risk, these investment opportunities fall under the category of "limited arbitrage".

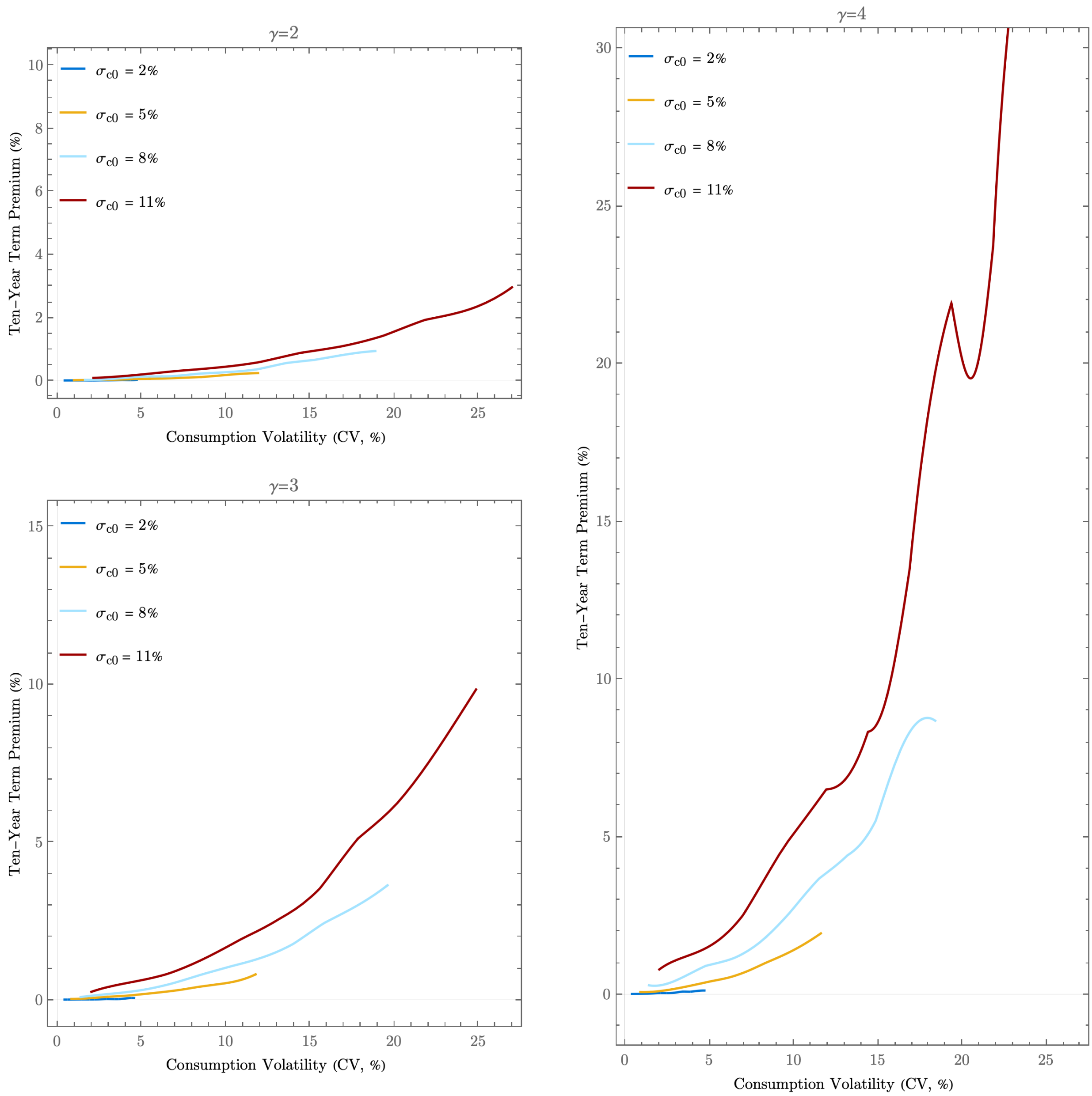


Figure 7: **Ten-year term premium in the TSU-HCV variation for different steady-state CV levels and risk aversion levels**
The plots show the ten-year term premium for different variations, and they are drawn with the same scale. Each plot corresponds to a different value of the risk aversion parameter γ, and each line corresponds to a different value for the steady state value of CV. The range of CV over which the lines are drawn, corresponds to the values of CV that can reasonably be acquired (these are the same ranges as in the previous figures).

investment opportunity brings more risk, and CV also rises. This setup can give rise to four separate variations depending on the behaviour of the short-term rate and the sign of ρ_{cx}. These are shown in Table 6. The movements in CD and CV have opposite effects on the short-term rate. Depending on the dominating component, the short-term rate can either increase or decrease in the magnitude of the investment opportunity. In addition, the sign of function A is fully determined by ρ_{cx}, which in turn depends on the portfolio composition of arbitrageurs, and how its value fluctuates given the changing state of the economy.[50] These two binary choices give rise to the four possibilities shown in Table 6.

Short-term rate	Positive ρ_{cx}	Negative ρ_{cx}
Short-term rate Decreasing with CV (CV dominates)	Arb-DP, **positive** term premia	Arb-DN, **negative** term premia
Short-term rate Increasing with CV (CD dominates)	Arb-IP, **negative** term premia	Arb-IN, **positive** term premia

Table 6: **Term premia sign in basic arbitrageur variations**

While each of these variations is plausible, the two that generate positive term premia are further analysed. In Arb-IN, term premia are positive and increasing with CV, as is the short-term rate. As shown in Figure 6, Arb-IN generates positive, time-varying and sizeable term premia. However, the size of the term premia is not as high as in TSU-Habit, TSU-HCV and Arb-DP. $\rho_{cx} < 0$ could be justified in Arb-IN, because an increase in the short-term rate could be inducing the arbitrageurs to invest more in the bond market and decrease their consumption. In addition, despite CD rising, an increase in the short-term rate can also imply a decrease in their wealth, if the arbitrageurs are bondholders.

In Arb-DP, $\rho_{cx} > 0$ can also be justified because it makes sense for consumption to increase when CD goes up. In addition, if the arbitrageurs are bondholders, then their wealth increases, as CV increases and the short rate goes down. This is also the variation that is most akin to the intuition provided in Vayanos and Vila (2021). As the short-term rate decreases, long-term bond yields underreact, and this leads to an increase in term premia. The arbitrageurs in Vayanos and Vila (2021) optimise between the mean and variance of their wealth, and consumption is not part of the analysis. To the best of my knowledge, this paper is the first to show that this behaviour can be rationalised within a consumption-based

[50]This is true to the extent that arbitrageurs do not have income external to their portfolio.

setup. Arb-DP also provides the characteristics of the consumption process that are consistent with term premia, and it shows that a low CV would not generate substantial variability in the arbitrageur setup. The mechanism driving term premia is basically the same as in TSU-HCV. Thus, explaining the main features of term premia requires a high CV. Concluding on whether actual bondholders' consumption process exhibits such volatility is not possible within this paper. However, the paper provides a theoretical prediction that can be evaluated and tested empirically. If such CV is judged to be too high, then arbitrageurs are likely not acting as consumers or on behalf of consumers. This would be evidence for the existence of frictions, such as the ones in the intermediary asset-pricing literature. Alternatively, if it is found that some bondholders have HCV as the model predicts, then it would be interesting to further research the reasons that distinguish these investors, and why they are not able to share their risk with the remaining population.

Apart from asset-pricing implications, the variations presented in this subsection are also significant for monetary policy, to the extent that monetary policy affects term premia (Beechey and Wright 2009). In particular, according to Arb-DP, central banks decreasing (increasing) interest rates is equivalent to increasing (decreasing) the CV of the marginal investors. An increasing CV implies higher term premia, and this mechanism hinges on stochastic consumption changes being positively correlated with CV. In addition, the effect on CV is very strong, as it can roughly range from 5% to 20%. On the contrary, the effect of monetary policy on the consumption process of non-investors might be muted, if they are indeed disconnected from the effects of bond markets. A full understanding of the effects of monetary policy on all agents in the economy would benefit from a full heterogeneous agent model that explains the investment behaviour of all households.[51]

Furthermore, the HCV and the arbitrageur variations have implications for household finance. In particular, the participants in these markets are assuming large consumption risks. Therefore, a usual household whose CV is low and whose utility function is similar to the utility function of the marginal investors could benefit from investing in long-term bonds, when term premia are high. This is valid, as long as CV of the household does not become too volatile due to this investment. However, the benefit is conditional on the state of the economy, and

[51] Schneider (2022) provided such a model, in which the state variable captures "aggregate conditions in the credit market". Similar models would be interesting, in which the state variable captures CD and CV.

it is not clear if the state of the economy is transparent to most households, as the current CV of marginal investors is not directly observable.[52] This advice would not be valid in the context of the habit model. In that case high term premia reflect states in which households have a high risk aversion, and investing in risky securities would not be appropriate.

6 Conclusion

In conclusion, standard models face three key challenges in explaining the features of term premia. Firstly, they typically generate long-term bonds that provide a hedge against risk, which leads to negative instead of positive term premia. Specifically, for a representative consumer, it is reasonable that a rise in CD is associated with a stochastic consumption increase. Therefore, bond prices increase when CD decreases, and vice versa. So, bonds are extra valuable, because they provide insurance against macroeconomic risk, and the associated term premia are negative. Similarly, for an aggregate representative consumer, it is reasonable that increased CV is associated with a stochastic consumption decrease. A similar argument implies that term premia are again negative. Secondly, time-varying CD generates constant instead of variable term premia. The paper shows that this turns out to be the case even in RU models. In contrast, time-varying CV always produces time-varying term premia, because by definition the state variable affects consumption uncertainty and, hence, risk. Thirdly, in calibrations according to an aggregate consumption process, term premia are typically very small in absolute value. The intuition for this is that consumption processes that are relatively stable give rise to term premia that are small. For term premia to be large it means that consumers are assuming large risks. Thus, given that aggregate consumption is relatively stable, the corresponding models imply low term premia. With the exception of the third shortcoming, these issues arise both in the TSU case and in the RU case.

The contribution of this paper is to generate realistic term premia in models that do not use time-varying risk aversion. In particular, model variations for which a) CV is high and stochastic, ranging for example from 5 to 20% per year, and for which b) stochastic consumption changes are positively correlated with

[52] One could argue that the state of the economy is directly observable by the level of the short-term rate. However, this paper has focused on explaining term premia, and a single state variable is used. In a full explanation of the dynamics of interest rates, at least two state variables would be needed. Hence, the level of the short-term rate would most likely not directly imply the level of term premia.

CV, can generate positive and highly variable term premia. The first component contributes to term premia being high in absolute value, and the second component implies that term premia are positive. Apart from time-varying risk aversion, this is also the *only* available consumption-based mechanism to generate positive and substantially time-varying term premia. An important implication of this model is the HCV for many states of the economy. The CV levels in these variations are similar or slightly lower to return volatility levels in the stock market. This implies two possibilities. The first is that marginal investors are a special group of actual consumers whose wealth is significantly volatile, possibly because it is composed of volatile assets like stocks; the second is that marginal investors correspond to intermediaries or other financial institutions whose value has this relatively high volatility. In order to answer or test this, it would be interesting to empirically investigate the specific investors that participate in the bond market, and measure their CV.

If term premia and by extension the real term structure of interest rates is primarily driven by a small group of investors, there are significant implications. Firstly, in the case that marginal investors correspond to financial intermediaries, this indicates that intermediaries are not investing according to the preferences or on behalf of the general public, whose consumption volatility is too low to justify these term premia. Secondly, it is critical for the conduct of monetary policy, because it implies that movements in term premia have a limited connection to the consumption of regular households. So, for instance, when the central bank tries to decrease long-term bond yields, this may be mostly affecting the economic situation of a small group of investors and not so much the economic situation of the general public,[53] whose consumption and purchasing behavior is normally the main point of interest of the monetary authority.

Further research could explain why marginal investors seem to be a small group. This can be done in a model that also accounts for households not participating in financial markets. Non-participation can be rationalised given the high volatility in financial markets and the existence of some friction. As a result, there would be reduced risk sharing, justifying CV being large. This setup is likely to jointly explain term premia, stock market non-participation, reduced risk sharing in the economy and the equity premium puzzle.

[53] Vayanos and Vila (2021), after a comment by John Cochrane, also make the point that monetary policy affecting the short-term rate can be viewed as a source of arbitrageur rent.

Appendix

A Definitions

In the following, a set of definitions is provided for the concepts used in the paper.[54]

- Throughout the paper, terms like yields, returns, term premia etc. should be understood as referring to their real counterparts, unless otherwise specified. The distinction is made explicit when necessary to avoid confusion.

- A **nominal zero-coupon bond** with maturity m is a security paying one unit of currency after m years.[55]

- A **real bond** with maturity m is a security paying one unit of currency times an adjustment, that corrects for the elapsed inflation from the time it was issued until its maturity. The payment occurs after m years. Equivalently, a real bond is a security that pays the value of some basket of goods[56] when it matures.[57]

- Q_t^m is the price of the bond with maturity m at time t.

- **Real (or nominal) yield** at time t of a real or (nominal) bond with maturity m years where Q_t^m is the price of the corresponding bond, which is perfectly liquid:[58]
$$y_t^m = \frac{-\log(Q_t^m)}{m}, \quad m > 0$$

- **Yield spread** at time t between maturity m and n, where typically $m > n$:
$$y_t^m - y_t^n$$

- **The yield curve or the term structure of interest rates** refers to yields as a function of maturity. The yield curve is sloping upward/downward (or the slope of the yield curve is positive/negative) when yields are an

[54]Including for some concepts which are mentioned in the main paper, without ever using in expressions.

[55]In the paper bonds always refer to zero-coupon bonds.

[56]Here there is an implicit assumption that individuals primarily care about this specific basket of goods. This basket of goods is also relevant for the calculation of inflation. Without this assumption, the study of real interest rates would be significantly hindered.

[57]A real bond of maturity $m+1$ one year ago is also equivalent to a real bond with maturity m today up to a renormalisation so that the principals match.

[58]Actual bonds' prices may deviate from Q_t^m due to liquidity considerations.

increasing/decreasing function of maturity. It is also possible that the slope is positive for some maturities and flat or negative for other maturities.

- **Annualised Gross Return** of a bond with maturity m from time t to $t+s$:

$$R^m_{t,t+s} = \sqrt[s]{Q^{m-s}_{t+s}/Q^m_t}$$

- **Log return or just return**[59] of a bond with maturity m from time t to $t+s$:

$$r^m_{t,t+s} = \log\left(R^m_{t,t+s}\right) = \frac{\log\left(Q^{m-s}_{t+s}\right) - \log\left(Q^m_t\right)}{s}$$

- **Instantaneous return** of a bond with maturity m at time t:

$$r^m_t = \lim_{s\to 0} r^m_{t,t+s}$$

- **Instantaneous short rate or just short rate** at time t:

$$r_t = \lim_{m\to 0} r^m_t = \lim_{m\to 0} y^m_t$$

- In the main paper yields are also referred to as *long-term interest rates*, whereas *interest rates* in general also include the short rate.

- **m-to-n year forward** at time t:

$$f^{m,n}_t = \frac{\log(Q^m_t) - \log(Q^n_t)}{n-m}$$

- **Instantaneous m-year forward** is:

$$f^m_t = \lim_{n\to m} f^{m,n}_t$$

- **Term or risk premium** of bond with maturity m at time t, where r_t is the instantaneous short-term rate of return at time t:[60]

$$TP^m_t = \frac{-\log(Q^m_t)}{m} - \frac{E_t\left[\int_0^m r_{t+s}ds\right]}{m}$$

  - If the term premium is zero for all m and t, this implies that the expected excess return from holding long-term bonds over any period is

[59] For convenience the term return is used to refer to log return.

[60] Equivalent definitions are given in discrete time by Cochrane and Piazzesi (2009).

also 0. This can be seen from the following equivalent definition, where rx_t^m is the instantaneous excess return from holding a bond of maturity m:[61]

$$TP_t^m = \frac{E_t\Big[\int_0^m r_{t+s}^{m-s} - r_{t+s} d\tau\Big]}{m} \equiv \frac{E_t\Big[\int_0^m rx_{t+s}^{m-s} d\tau\Big]}{m}$$

– Here the following is used:

$$\begin{aligned}
-\log(Q_t^m) =& \Big(-\log\left(Q_t^m\right) + \log\left(Q_{t+m/N}^{m-m/N}\right)\Big) + \Big(-\log\left(Q_{t+m/N}^{m-m/N}\right) + \log\left(Q_{t+2m/N}^{m-2m/N}\right)\Big) + \\
& \ldots + \Big(-\log\left(Q_{t+m-m/N}^{m/N}\right) + \underbrace{\log\left(Q_{t+m}^{0}\right)}_{=0}\Big) \\
=& \frac{m}{N}\left(r_{t,t+m/N}^{m} + r_{t+m/N,t+2m/N}^{m-m/N} + \ldots + r_{t+m-m/N,t+m}^{m/N}\right) \\
\approx& \int_0^m r_{t+s}^{m-s} ds \qquad (24)
\end{aligned}$$

where N is some positive integer. The last line follows by N going to infinity, which means that the sum becomes an integral and the returns become instantaneous returns.

– Since Q represents the price of a bond that is perfectly liquid, the term premium does not include a liquidity premium.

- The quantity used above is also referred to:

$$\frac{E_t\Big[\int_0^m r_{t+s} ds\Big]}{m}$$

as **risk-neutral yield** of bond with maturity m.

- **Term or risk premium** of m-to-n year maturity forward at time t, where r_t is the instantaneous short-term rate of return at time t:

$$TP_t^{m,n} = \frac{\log(Q_t^m) - \log(Q_t^n)}{n-m} - \frac{E_t\Big[\int_{t+m}^{t+n} r_\tau d\tau\Big]}{n-m}$$

- The second term on the right-hand side of the equation above is the risk-neutral m-to-n year forward.

[61] If the excess return were positive for any period, then the expected term premium for the remaining period would be negative. This violated the initial assumption.

- In the paper many of the variables introduced here depend on time only through the state variable. So they will be denoted instead as:

$$Q(x_t, m), r(x_t), TP(x_t, m)$$

- The main paper also refers to the **value of the risk-neutral bond**. This is the implied value attached to a bond by a risk-neutral investor and it can be defined based on the risk-neutral yield defined above:[62]

$$H(x_t, m) = e^{-E_t\left[\int_0^m r_{t+s} ds\right]}$$

- The **strong version of the Expectations Hypothesis** holds when:

$$TP_t^m = 0, \quad \text{for all } m$$

- The **weak version of the Expectations Hypothesis** holds when:

$$TP_t^m = g(m), \quad \text{for all } m$$

  where g is some function of maturity, independent of the state of the economy and independent of time.

- **Predictability** refers to the ability to predict movements in excess returns. The prediction could be based on any information, but the literature has focused on using information in yields to predict subsequent yields in the future.

- **Excess volatility** of interest rates refers to long-term interest rate variations that are too large to be explained by the variation of the short rate alone while keeping the discount rate constant.[63]

[62] In the main paper, in Section 4.7.3, equivalent definition , which also shows the intuition regarding the calculation of the term premium in this paper.

[63] To be completely precise excess volatility needs to be defined in terms of some benchmark model. As excess volatility is not directly investigated, it is not defined explicitly.

B Explanatory power of the principal components of real interest rates

Apart from Figure 2, a series of regressions is also performed to demonstrate the strong dependence of nominal rates on real rates. In particular, the first two principal components are extracted from a series of real yields with different maturities.[64] Only two components are used because they explain more than 99.95% of the variance of real yields. Next, nominal yields and nominal yield spreads are regressed on these two principal components.[65] Indeed, the result is that the information contained within real rates explains most of the movements of nominal rates. The results are shown in Table 7. The coefficients are highly significant for both components, but more importantly, the R-squared is high in these regressions. For the level regressions, it ranges from 87% to 93%, while for the spread regressions, it ranges from 69% to 79%. Thus, both the level and the spread of nominal rates are mostly explained by the information and hence the processes that generate the real term structure.

Table 7: Regressions of the level and the spread of nominal bonds on the principal components extracted from the real term structure

	5 yr	10 yr	5-10 yr spread	15 yr	5-15 yr spread	20 yr	5-20 yr spread
Intercept	2.94***	3.73***	0.79***	4.13***	1.19***	4.29***	1.35***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
comp1	0.28***	0.26***	-0.02***	0.25***	-0.04***	0.23***	-0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
comp2	0.43***	-0.24***	-0.66***	-0.55***	-0.97***	-0.67***	-1.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R-squared	0.87	0.93	0.69	0.93	0.74	0.93	0.79
R-squared Adj.	0.87	0.93	0.69	0.93	0.74	0.93	0.79

[64] The principal components are extracted from yields of all yearly maturities from two to twenty years.

[65] A similar exercise in performed by Abrahams *et al.* (2016) and they also find similar results. In their case, it is the real rates that are regressed on the principal components of the nominal rates. The inverse exercise is performed because the focus here is how much nominal rates are explained by real rates.

C Components of the pricing equation

This section provides an explanation for each part of the pricing equation (16) which is repeated here:

$$-Q_m - r(x_t)Q + \big(\log(\phi)x_t + A(x_t)\big)Q_x + \frac{\sigma_{xt}^2}{2}Q_{xx} = 0 \tag{25}$$

- In the simplest case $\phi = 1$, $A(x_t) = 0$ and $\sigma_x(x_t) = 0$ for all x_t. Then the equation is:

$$Q_m = -r(x_t)Q = -x_tQ$$

  This corresponds to an economy with a constant state. Figure 8 shows that in this economy yields are always equal to the short rate, term premia are equal to 0 and given the state of the economy nothing will ever change.

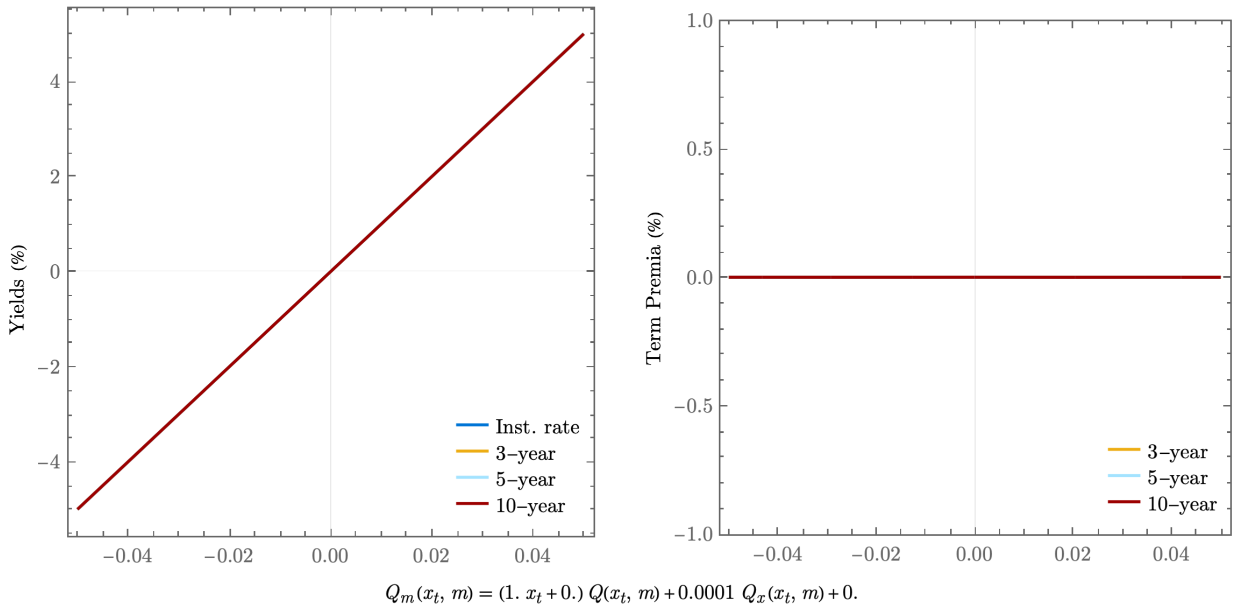


Figure 8: The left plot shows the short-term rate and yields of different maturities as a function of the state variable. The right plot shows the term premia for different maturities as a function of the state variable.

- $\phi \neq 1$:

$$Q_m = r(x_t)Q - \log(\phi)x_tQ_x = x_tQ - \log(0.9)x_tQ_x$$

  Here there is again no volatility of the state variable. Thus, this corresponds to a deterministic economy. However, the state is not constant, it drifts towards the state $x_t = 0$, which can be thought of as the steady state. This implies that long-term yields will lie between the contemporaneous short rate and the steady-state short rate. As shown in Figure 9 this results in a characteristic picture, in which all yields intersect at the steady state. If the process moved towards the steady state faster (lower ϕ), then the yields would be more spread out. Given that there is no uncertainty, the term premia are again zero.

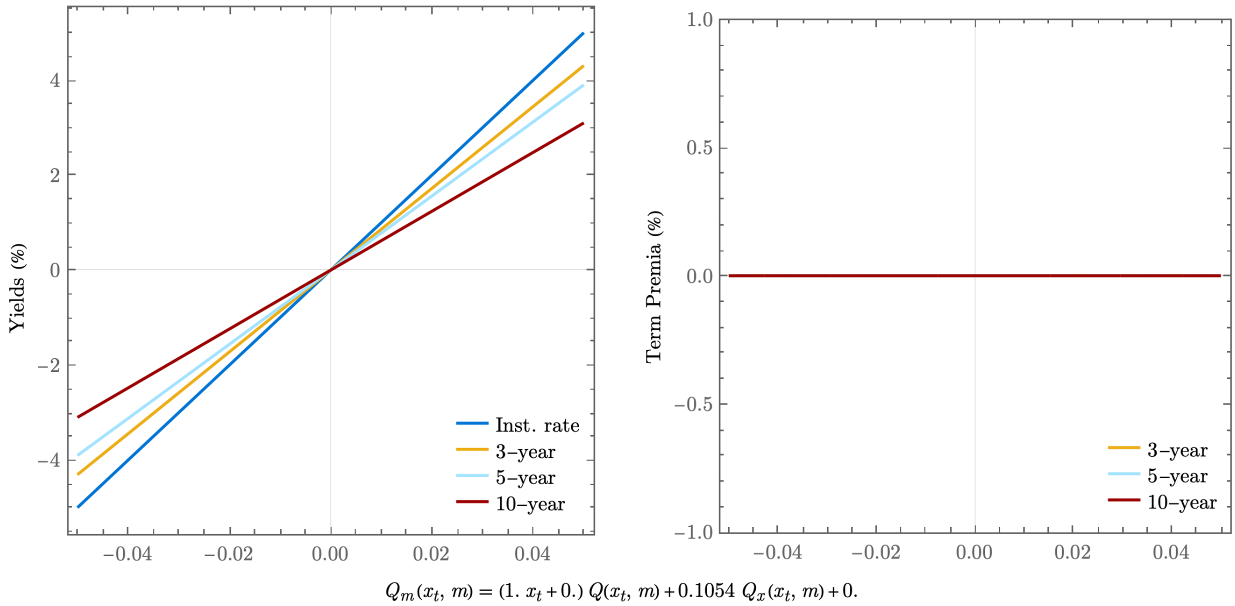


Figure 9: The left plots show the short-term rate, the five-year yield and the five-year risk-neutral yield as a function of CD. The right plot shows the term premia for different maturities as a function of the state variable.

- $A(x_t) = c \neq 0$:

$$Q_m = r(x_t)Q - \big(\log(\phi)x_t + A(x_t)\big)Q_x = x_tQ - (\log(0.9)x_t + 0.01)Q_x$$

As stated in the main paper A generates term premia. This case does not directly correspond to some economic situation because, the state variable volatility is again 0, and in the actual economic models this also implies $A(x_t) = 0$. However, for intuition, the resulting "yields" and "term premia" are shown. As Figure 10 shows, now the yields do not intersect at the steady state. Now the longer-term yields are higher at the steady state. This implies positive term premia and indeed as shown in the right panel, term premia are positive, proportional to the maturity of the bond and constant as a function of the state variable. The latter fact is due to $A(x_t)$ being constant for all x_t and the fact that yields are linear. Finally, the term premia are positive, because A is positive and the short rate is increasing with respect to the state variable.

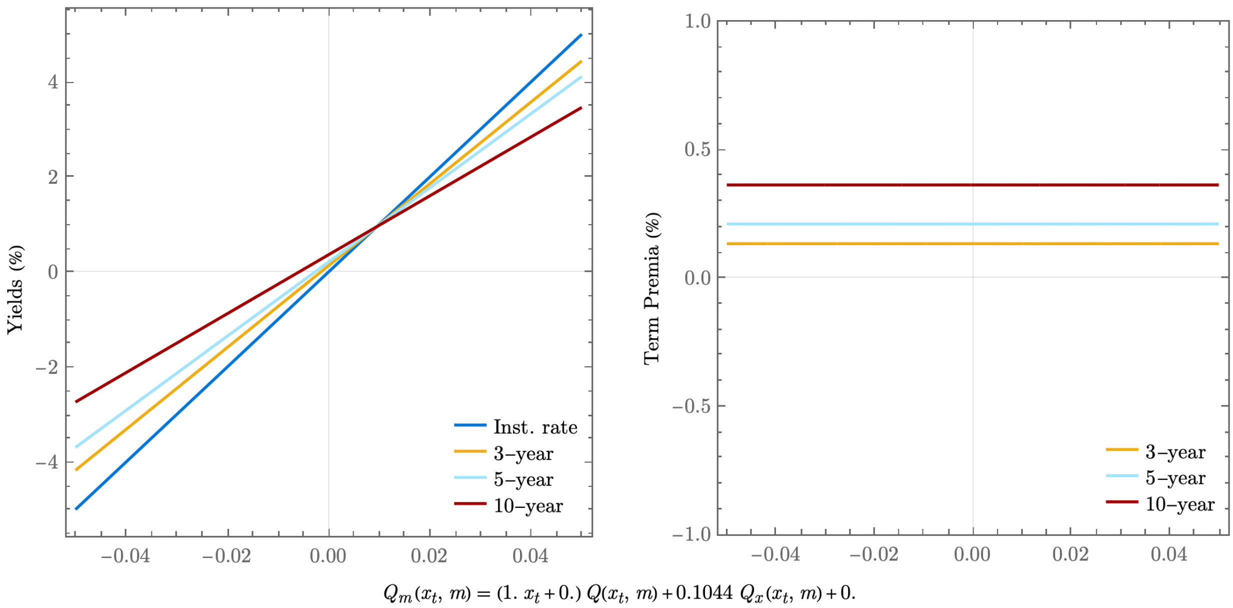


Figure 10: The left plots show the short-term rate, the five-year yield and the five-year risk-neutral yield as a function of CD. The right plot shows the term premia for different maturities as a function of the state variable.

- $A(x_t) = 0.0005 + 0.02x_t$. This means that now A changes with the state variable. The result is shown in Figure 11. Term premia follow the behaviour of A. The correspondence would not be so close if the short rate were a non-linear function of the state variable.

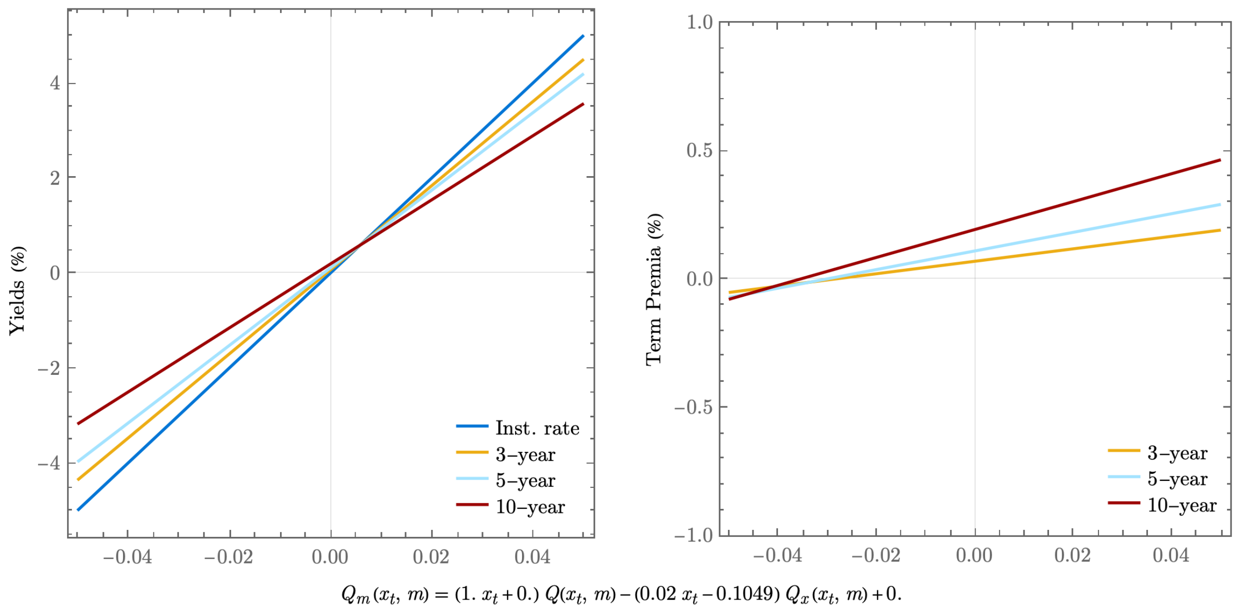


Figure 11: The left plots show the short-term rate, the five-year yield and the five-year risk-neutral yield as a function of CD. The right plot shows the term premia for different maturities as a function of the state variable.

- $\sigma_{xt} \neq 0$:

$$Q_m = r(x_t)Q - \log(\phi)x_tQ_x + \frac{\sigma_{xt}^2}{2}Q_{xx} = x_tQ - \log(0.9)x_tQ_x - \frac{0.03^2}{2}Q_{xx}$$

  Here $A(x_t) = 0$. Thus, the effect of volatility can be seen. This case corresponds to a case where there is volatility of the short rate, but there is again no priced risk. So there is no risk premium. This can be seen on the right panel of Figure 12.[66] Nevertheless, the yields are not the same as in the deterministic case with steady-state reversion, as they do not intersect at the steady state. The long-term yields are pushed downwards, and, even though it might not be obvious, the effect of uncertainty increases more than linearly with maturity. This effect is due to the so-called convexity that is common in finance. In particular, the price of the long-term bond is a decreasing convex function of the short rate and this implies that lower interest rates have a higher effect on the price of the bond, especially for long maturities. Thus, given that there is variation and a chance for the short rate to reach lower levels, these will outweigh the high rates, and push long-term yields

[66] The term premia do not look completely flat because the Monte-Carlo calculation has some uncertainty in the calculation.

downward. Finally, this also means that a downward-sloping term structure does not necessarily imply negative term premia.

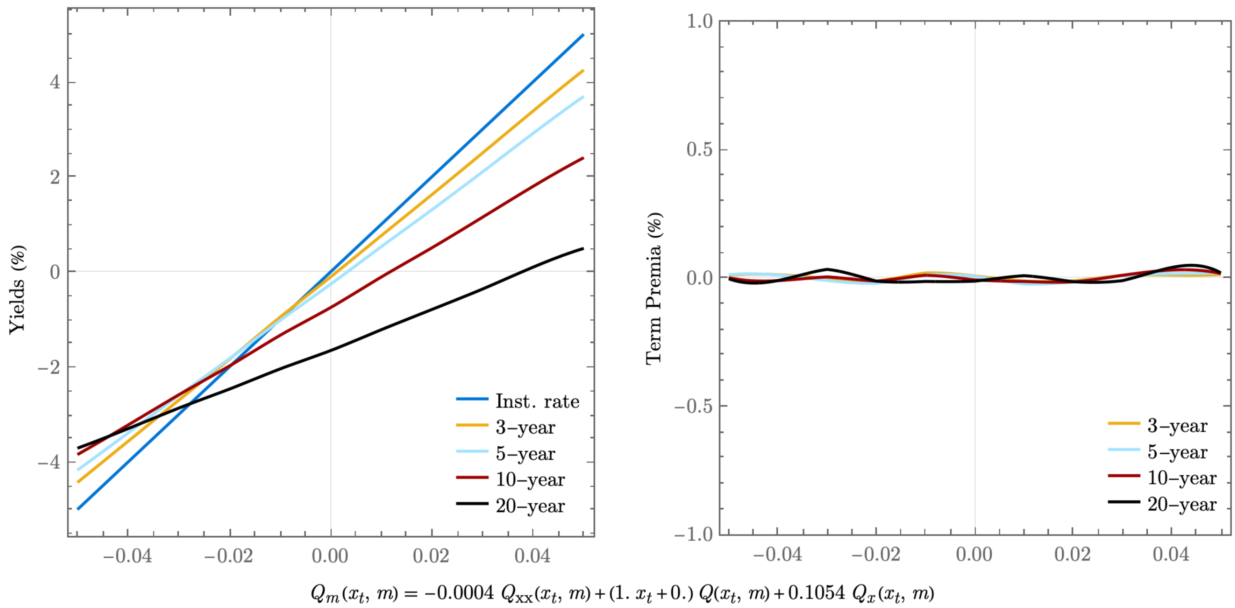


Figure 12: The left plots show the short-term rate, the five-year yield and the five-year risk-neutral yield as a function of considers. The right plot shows the term premia for different maturities as a function of the state variable.

- full case:

$$Q_m = r(x_t)Q - \log(\phi)x_tQ_x + \frac{\sigma_{xt}^2}{2}Q_{xx} = x_tQ - \big(\log(0.9)x_t + 0.001\big)Q_x - \frac{0.005^2}{2}Q_{xx}$$

This case contains all the components. Unlike the previous case, as can be seen in Figure 13, the yields seem to intersect close to the steady state. Thus, the yield curve would often be flat in this economy. However, term premia are positive. The yields are close to flat at the steady state because term premia and convexity largely cancel each other out.

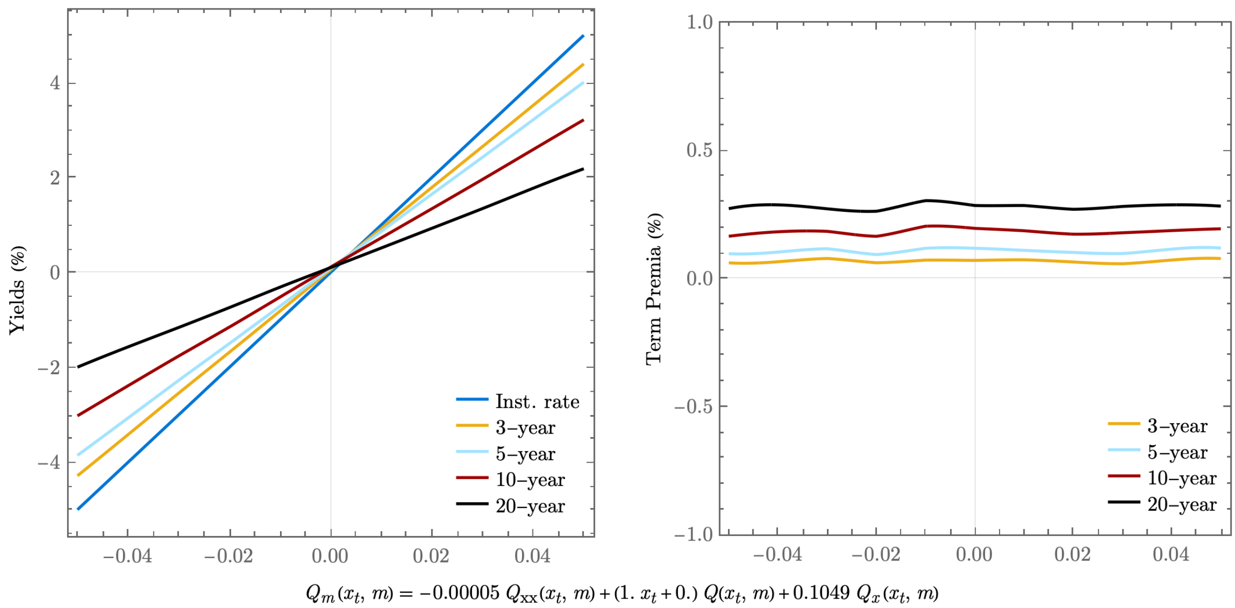


Figure 13: The left plots show the short-term rate, the five-year yield and the five-year risk-neutral yield as a function of CD. The right plot shows the term premia for different maturities as a function of the state variable.

D Calibration of the state variable volatility

As mentioned in Subsection 5.1, the paper aims to simultaneously match the variability of term premia and the variability of the short rate. This is achieved by calculating the range of the two-year TIPS security over the available sample in the Gürkaynak *et al.* (2010) dataset.[67] The result is a range of 7.27%.[68] Time-series are then simulated with twelve-year duration[69] for all the variations that are investigated. Based on these simulations the range sizes are ranked and the tenth quantile is made equal to the range in the data. This is done for the models that are not able to produce highly variable term premia, to give these models the benefit of the doubt and the best chance to succeed. Namely, it is possible that the observed short rate volatility has been by chance relatively low and the underlying process is significantly more volatile. Thus, the model variations are made as volatile as possible to generate as large a time variability in term premia

[67] Two years is the shortest maturity in the data.

[68] This could be overestimating the plausible range as the maximum was achieved during the financial crisis when the TIPS market was not behaving normally.

[69] This matches the length of the sample in Abrahams *et al.* (2016), but this should arguably be changed to match the length of the sample in Gürkaynak *et al.* (2010). In any case, the length of that sample is approximately 15 years.

as possible. For the models that succeed in producing significantly time-varying term premia, it is again verified that the empirical volatility, as expressed by the observed range, falls within the model predictions. For each model variation, the figures in Appendix F display the value of the empirical range and the values of the model-implied tenth and ninetieth quantile ranges.

E Term premia measures

E.1 Figure from Abrahams et al. (2016)

TIPS Term Premium

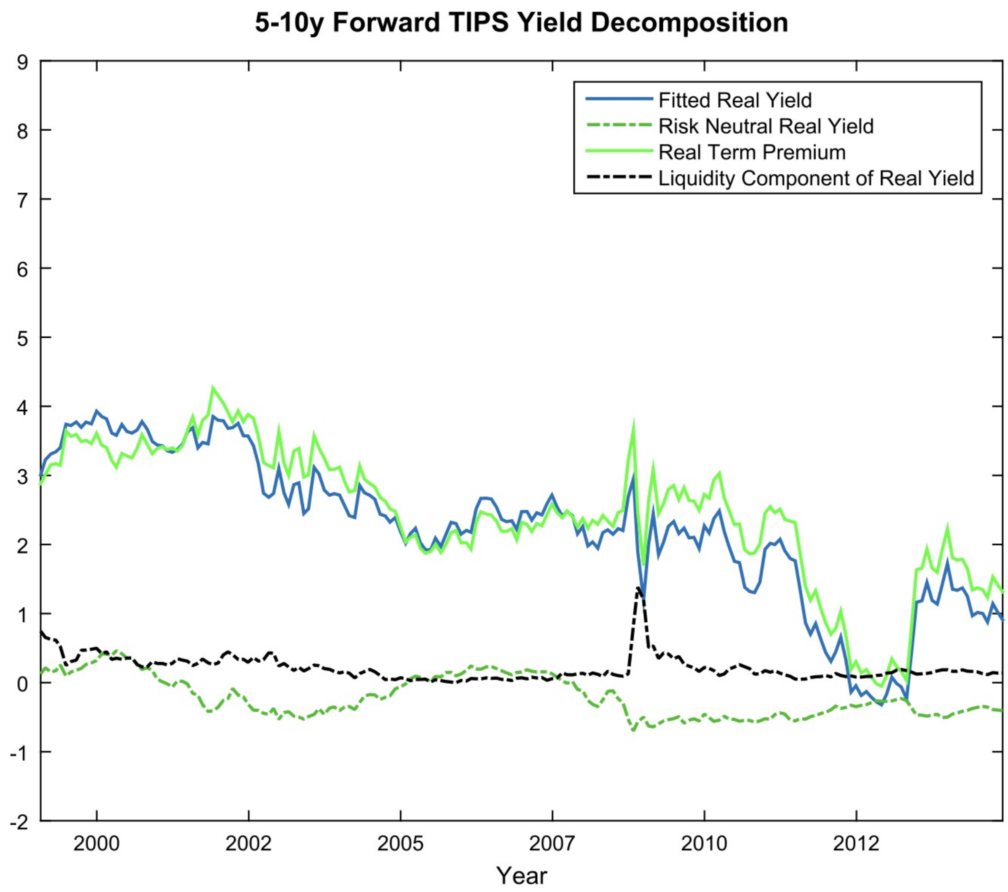


Figure 14: The figure shows the time series of the five-to-ten-year forward term premium along with its decomposition to the risk-neutral yield, the term premium and the liquidity premium. The same decomposition in the figures in this paper are also shown as a function of the state variable.

E.2 Term premium based on d' Amico et al. (2018)

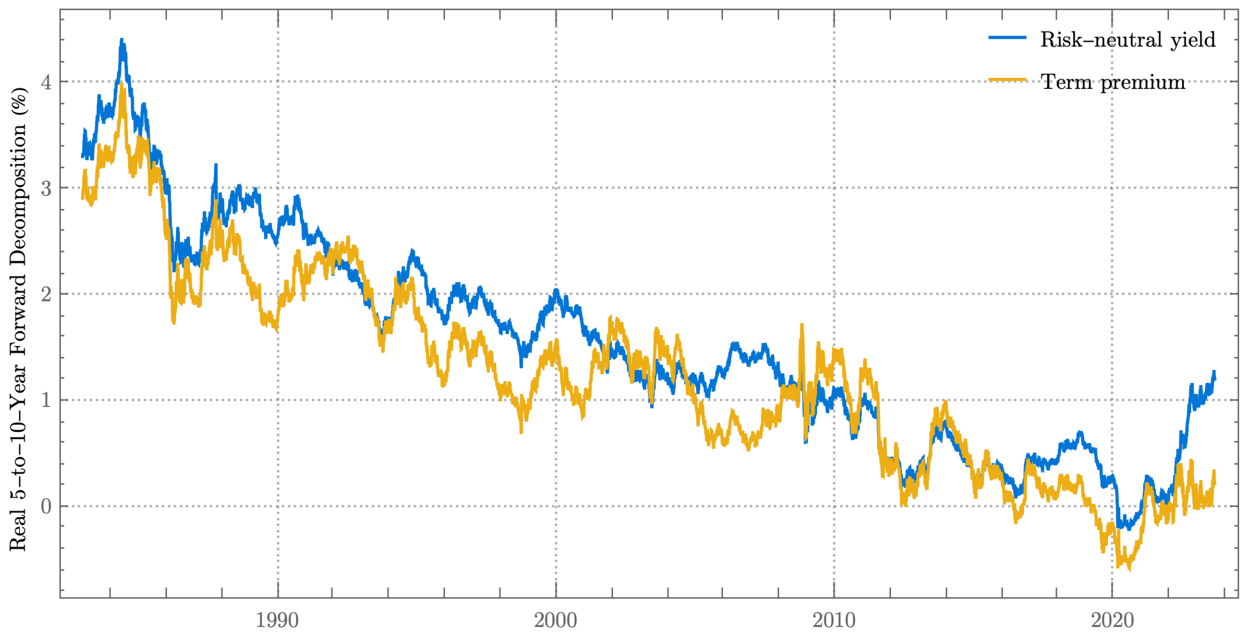


Figure 15: Time series of the forward 5-to-10-year term premium for the US. This is the same quantity as the solid green line in Figure E.1 from Abrahams *et al.* (2016).
Data Source:
https://www.federalreserve.gov/econres/notes/feds-notes/tips-from-tips-update-and-discussions-20190521.html

F Yields and term premia in other model variations – Time-separable utility

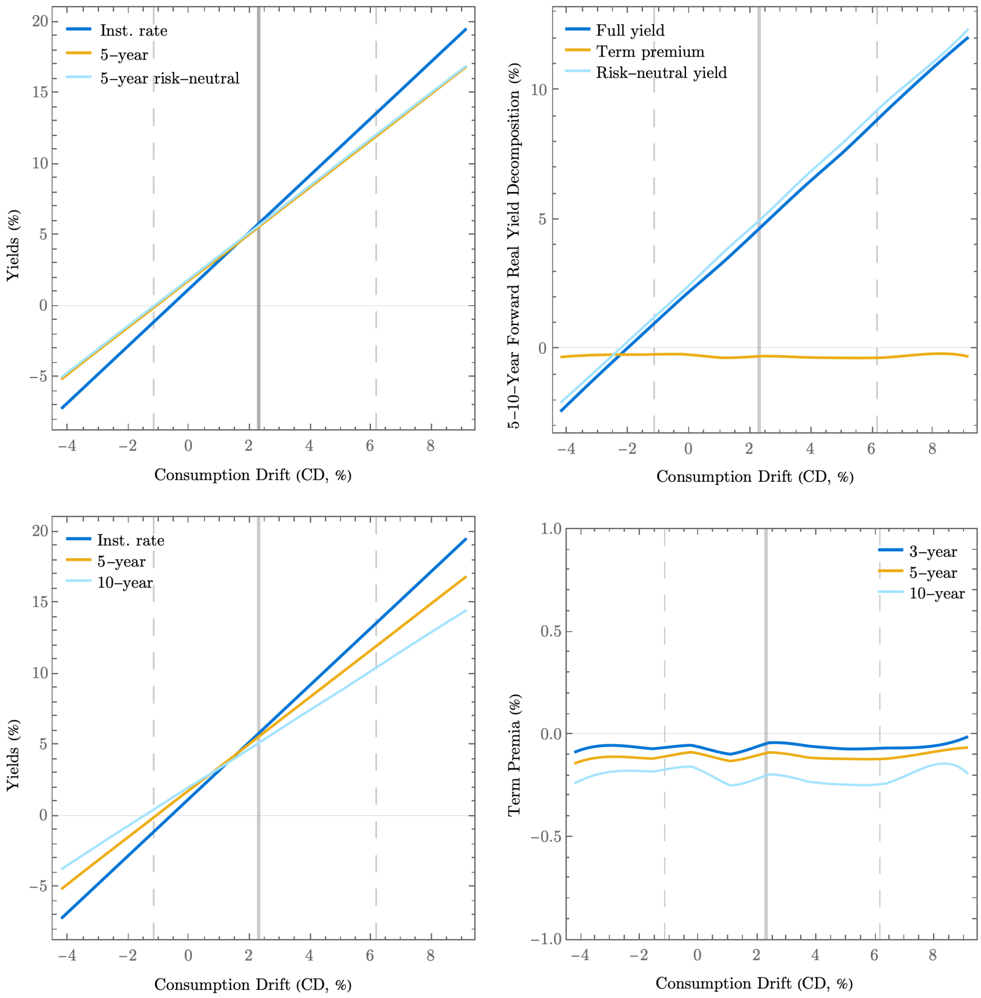
In this part, more plots are presented for the variations discussed in the main paper, and results for other model variations are also presented. These other variations should reinforce the conclusions in the main paper as a long series of calibrations is examined. The upper left and upper right plots are the same as in the main paper. The lower left plot shows the level of yields for different maturities as a function of the state variable. The lower right plot shows the level of the term premium for different maturities as a function of the state variable. Again each figure states the exact specification.

F.1 Model index

Names of all model variations are shown in Appendix F. The abbreviations used here are time-varying (TV), consumption drift (CD), consumption volatility (CV), time-separable utility (TSU), recursive utility (RU), intertemporal elasticity of substitution (IES).

Model Variation Description	**Abbreviation**	**References**
TV CD with TSU.	TSU-CD	Figure 16
TV CD with TSU and high risk aversion.	TSU-CD-HRA	Figure 17
TV CD with TSU and low persistence.	TSU-CD-LP	Figure 18
TV CD with TSU and high correlation ρ_{cx}.	TSU-CD-HCor	Figure 19
TV CD with TSU and high impatience.	TSU-CD-HImp	Figure 20
TV and high CD with TSU.	TSU-HCD	Figure 21
TV CD with TSU and high CV.	TSU-CD-HCV	Figure 22
TV CV with TSU.	TSU-CV	Figure 23
TV CV with TSUand high risk aversion.	TSU-CV-HRA	Figure 24
TV CV with TSU and high CD.	TSU-CV-HCD	Figure 25
TV and HCV with TSU and positive correlation ρ_{cx}.	TSU-HCV	Figure 26
TV and HCV with TSU and negative correlation ρ_{cx}.	TSU-HCV-NCor	Figure 27
Both TV CD and CV, short-term rate decreasing in CV and ρ_{cx} positive.	TSU-Arb-DP	Figure 28
Both TV CD and CV, short-term rate increasing in CV and ρ_{cx} negative.	TSU-Arb-IN	Figure 29
Both TV CD and CV, short-term rate decreasing in CV and ρ_{cx} negative.	TSU-Arb-DN	Figure 30
Both TV CD and CV, short-term rate increasing in CV and ρ_{cx} positive.	TSU-Arb-IP	Figure 31
TV external habit with TSU.	TSU-Habit	Figure 32
TV external habit with TSU and low b.	TSU-Habit-Low.b	Figure 33
TV external habit with TSU and $b < 0$.	TSU-Habit-Neg.b	Figure 34
TV external habit with TSU with constant state variable volatility.	TSU-Habit-CSV	Figure 35
TV CD with RU.	RU-CD	Figure 36
TV CD with RU and high risk aversion.	RU-CD-HRA	Figure 37
TV CD with RU with high IES.	RU-CD-HIES	Figure 38
TV CD with RU with Low IES.	RU-CD-LIES	Figure 39
TV CD with RU with high ρ_{cx}.	RU-CD-HCor	Figure 40
TV CD with RU with ρ_{cx} negative.	RU-CD-NCor	Figure 41
TV and high CD with RU.	RU-HCD	Figure 42
TV CD with RU and high CV.	RU-CD-HCV	Figure 43
TV and heteroskedastic CD with RU and ρ_{cx} positive.	RU-CD-Heterosk-PCor	Figure 44
TV and heteroskedastic CD with RU and ρ_{cx} negative.	RU-CD-Heterosk-NCor	Figure 45
TV CV with RU.	RU-CV	Figure 46
TV CV with RU with high risk aversion.	RU-CV-HRA	Figure 47
TV CV with RU and high persistence IES.	RU-CV-HP	Figure 48
TV CV with RU and high IES.	RU-CV-HIES	Figure 49
TV CV with RU and low IES.	RU-CV-LIES	Figure 50
TV and HCV with RU and ρ_{cx} positive.	RU-HCV-PCor	Figure 51
TV and HCV with RU and ρ_{cx} negative.	RU-HCV-NCor	Figure 52

F.2 TSU-CD, calibration used in the main paper, Figure 3



γ=2, ρ=0.012, $\mu_{ct}=x_t+0.0252$, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5

$\mathrm{range}_{low}(r_{2,model})$=0.10958, $\mathrm{range}_{high}(r_{2,model})$=0.202976, $\mathrm{range}(r_{2,data})$=0.071748.

Figure 16: Time-varying CD with TSU.
The left plot shows the short-term rate, the five-year yield and the five-year risk-neutral yield as a function of consumption growth. The right plot shows the decomposition of the five-to-ten-year forward into the term premium and the risk-neutral components. The solid vertical line shows the level of the ergodic median, and the left and right dashed vertical lines show the median minimum and maximum value respectively over a series of simulations for 12 years. This means that half the simulated paths were below the right dashed line and half the simulated paths were above the left dashed line. The left and right boundaries are the 10th percentile of minimum values and the 90th percentile of maximum values from the same simulations. This means that 90% of simulated paths were above the left boundary and 90% of simulated paths were below the right boundary.

F.3 TSU-CD-HRA, $\gamma = 8$

Term premia are a bit larger, but again negative and constant with respect to the state variable.

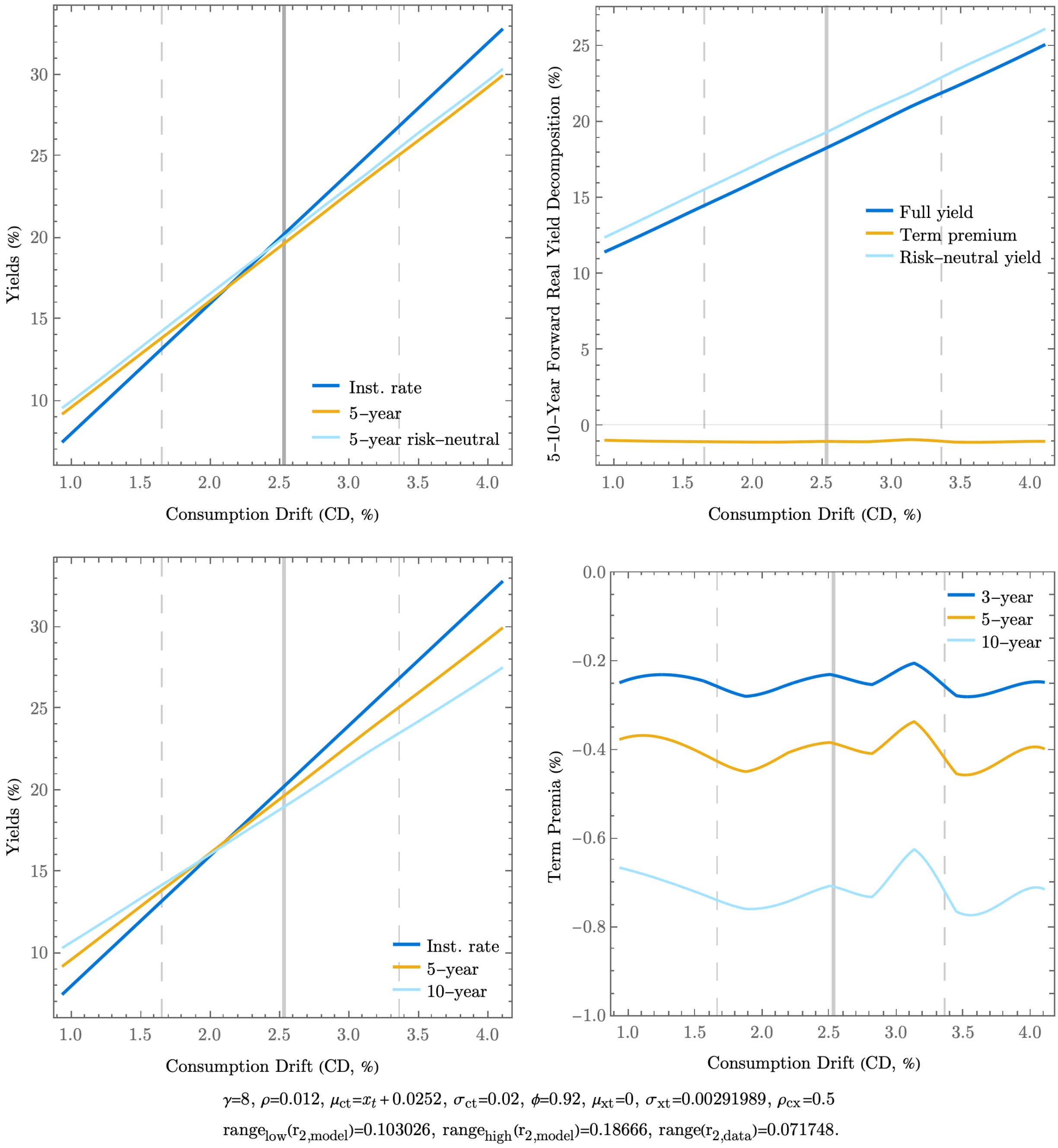


Figure 17: Time-varying CD with TSU and higher risk aversion.
See Figure 16 for more details about the plots.

(variation overview)

F.4 TSU-CD-LP, $\phi = 0.8$

Nothing changed in the term premia. There is a larger separation between yields similar to the corresponding mechanism in Appendix C.

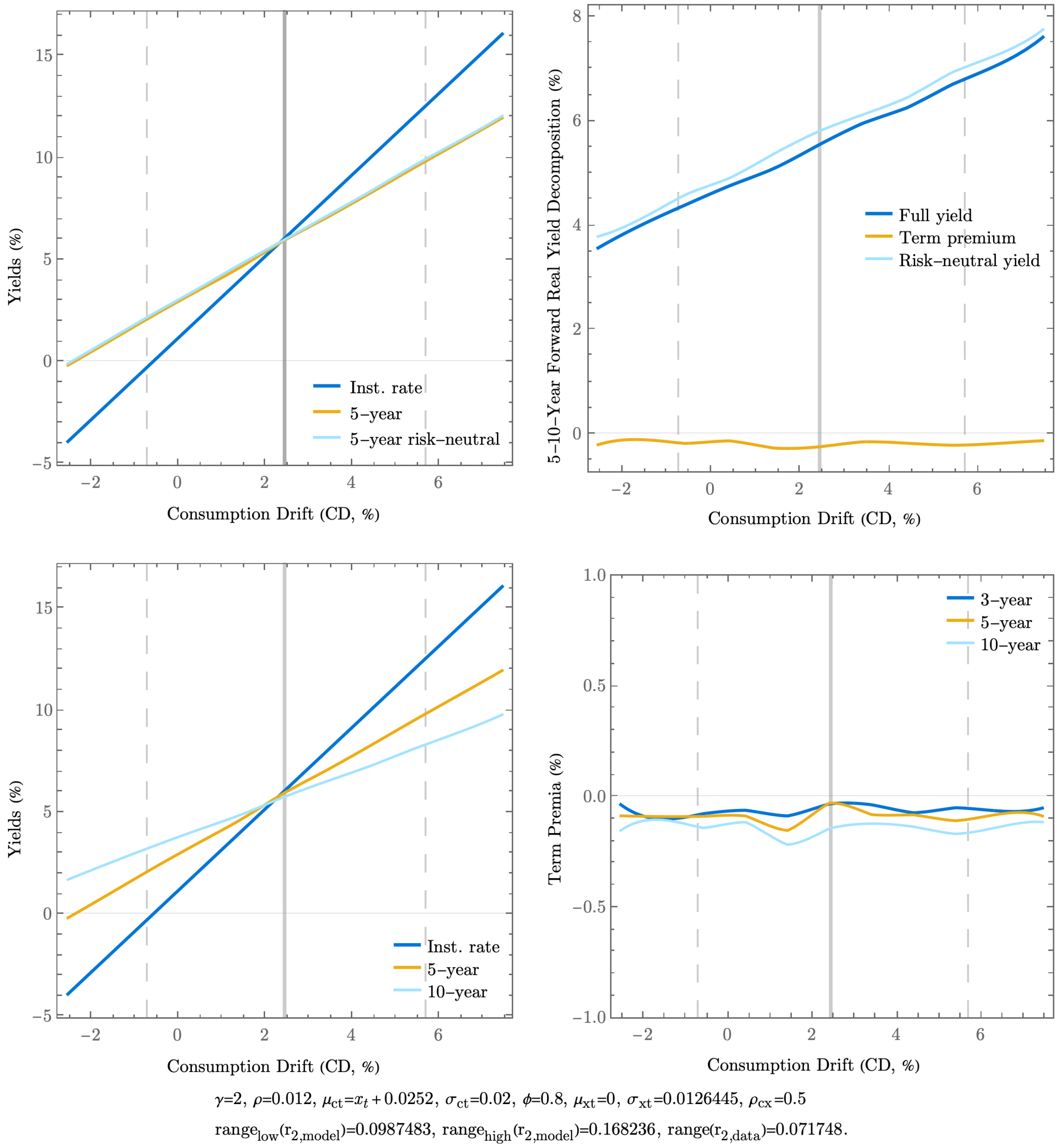
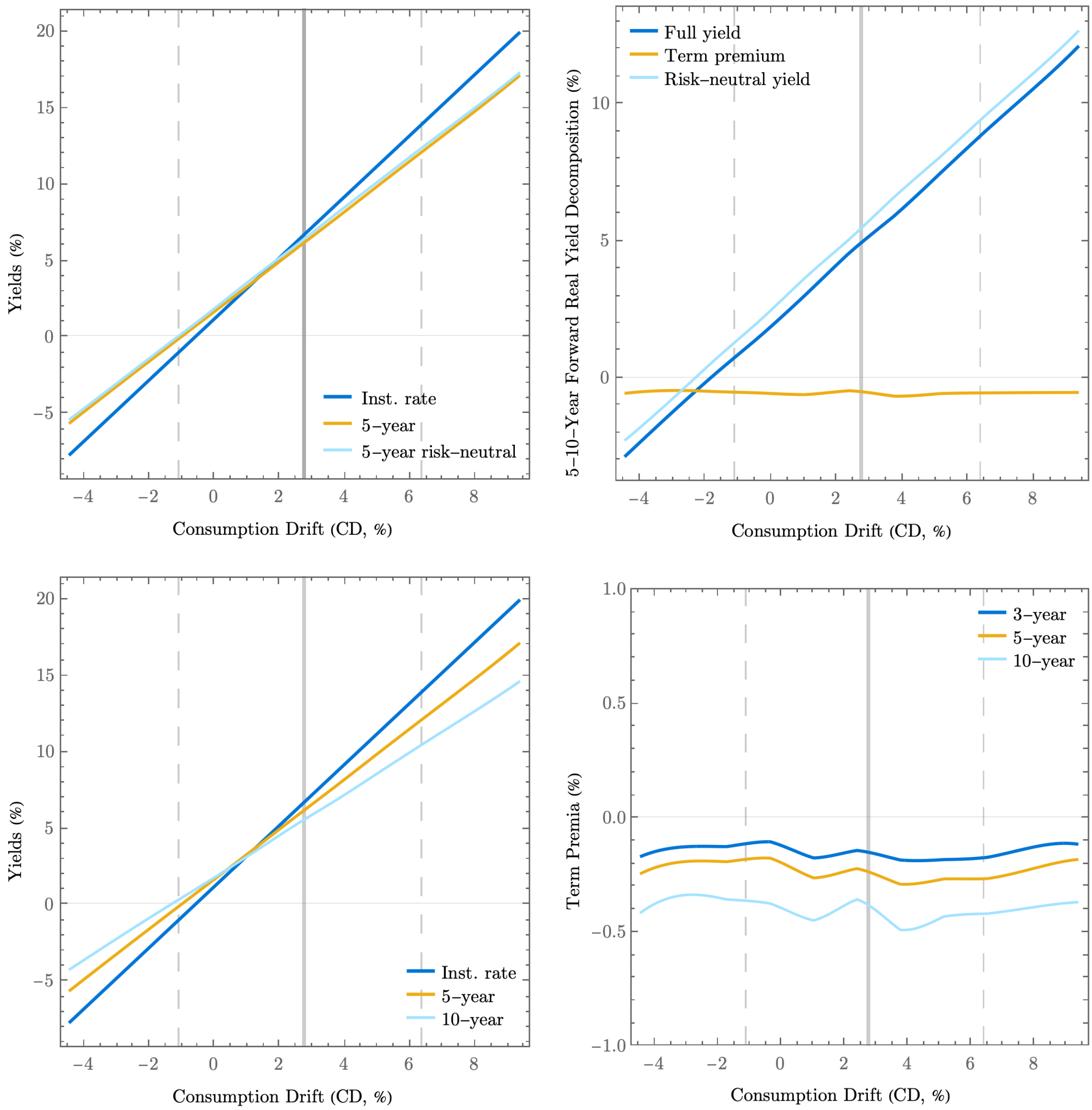


Figure 18: Time-varying CD with TSU and low persistence.
See Figure 16 for more details about the plots.

(variation overview)

F.5 TSU-CD-HCor, $\rho_{cx} = 1$

The term premia are larger in absolute value.



γ=2, ρ=0.012, μ_{ct}=x_t + 0.0252, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=1

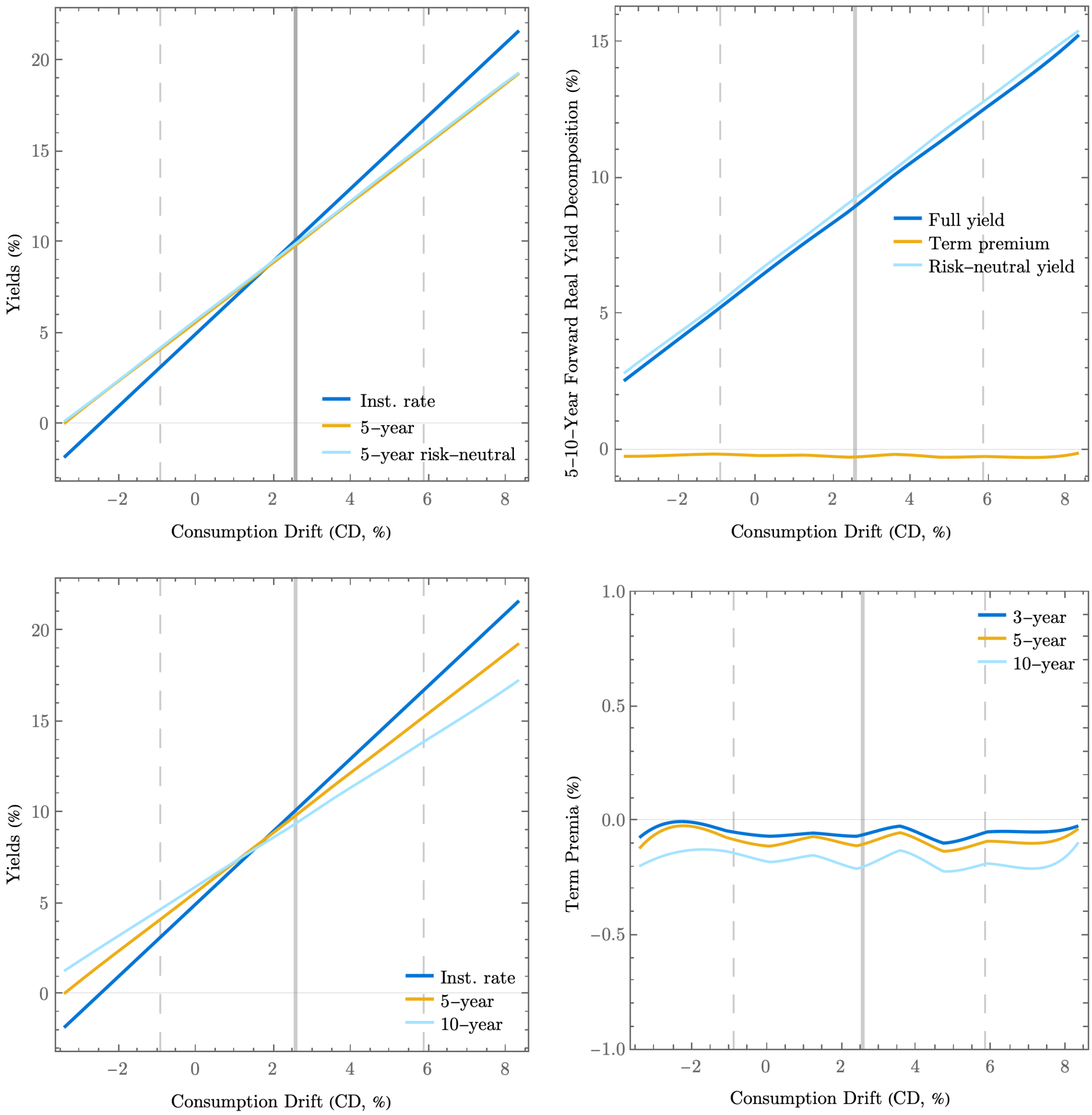
$\text{range}_{\text{low}}(\text{r}_{2,\text{model}})$=0.111377, $\text{range}_{\text{high}}(\text{r}_{2,\text{model}})$=0.203865, $\text{range}(\text{r}_{2,\text{data}})$=0.071748.

Figure 19: Time-varying CD with TSU and high correlation ρ_{cx}.
See Figure 16 for more details about the plots.

(variation overview)

F.6 TSU-CD-HImp, $\rho = 0.05$

Yields move higher without any change in term premia.



γ=2, ρ=0.05, $\mu_{\rm ct}$=x_t + 0.0252, $\sigma_{\rm ct}$=0.02, ϕ=0.92, $\mu_{\rm xt}$=0, $\sigma_{\rm xt}$=0.011495, $\rho_{\rm cx}$=0.5

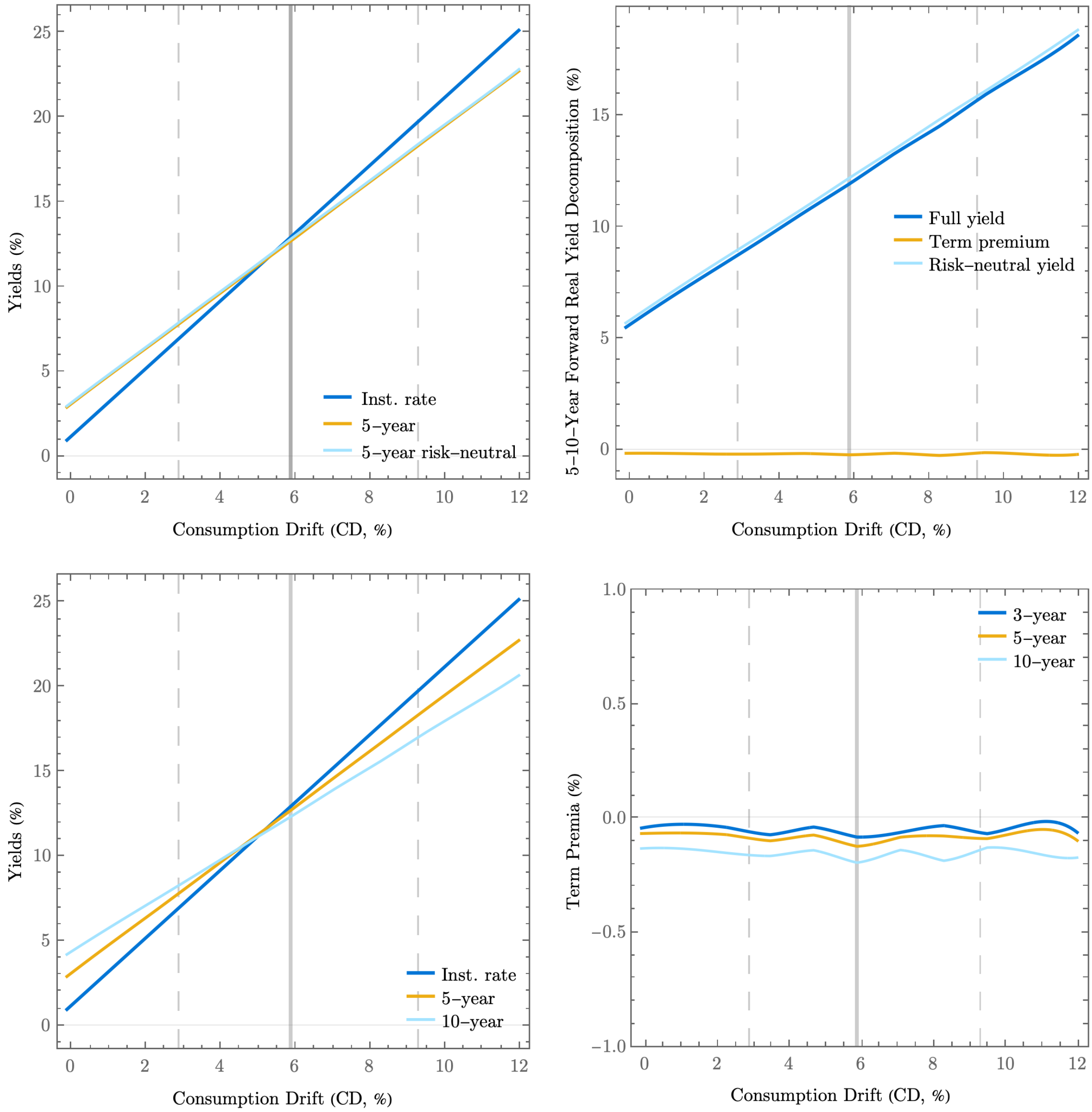
range$_{\rm low}$(r$_{2,\rm model}$)=0.0973497, range$_{\rm high}$(r$_{2,\rm model}$)=0.182416, range(r$_{2,\rm data}$)=0.071748.

Figure 20: Time-varying CD with TSU and high impatience.
See Figure 16 for more details about the plots.

(variation overview)

F.7 TSU-HCD, $\mu_{c0} = 0.06$

Again, yields move higher without any change in term premia.



γ=2, ρ=0.012, $\mu_{\text{ct}}=x_t+0.06$, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.011495, ρ_{cx}=0.5

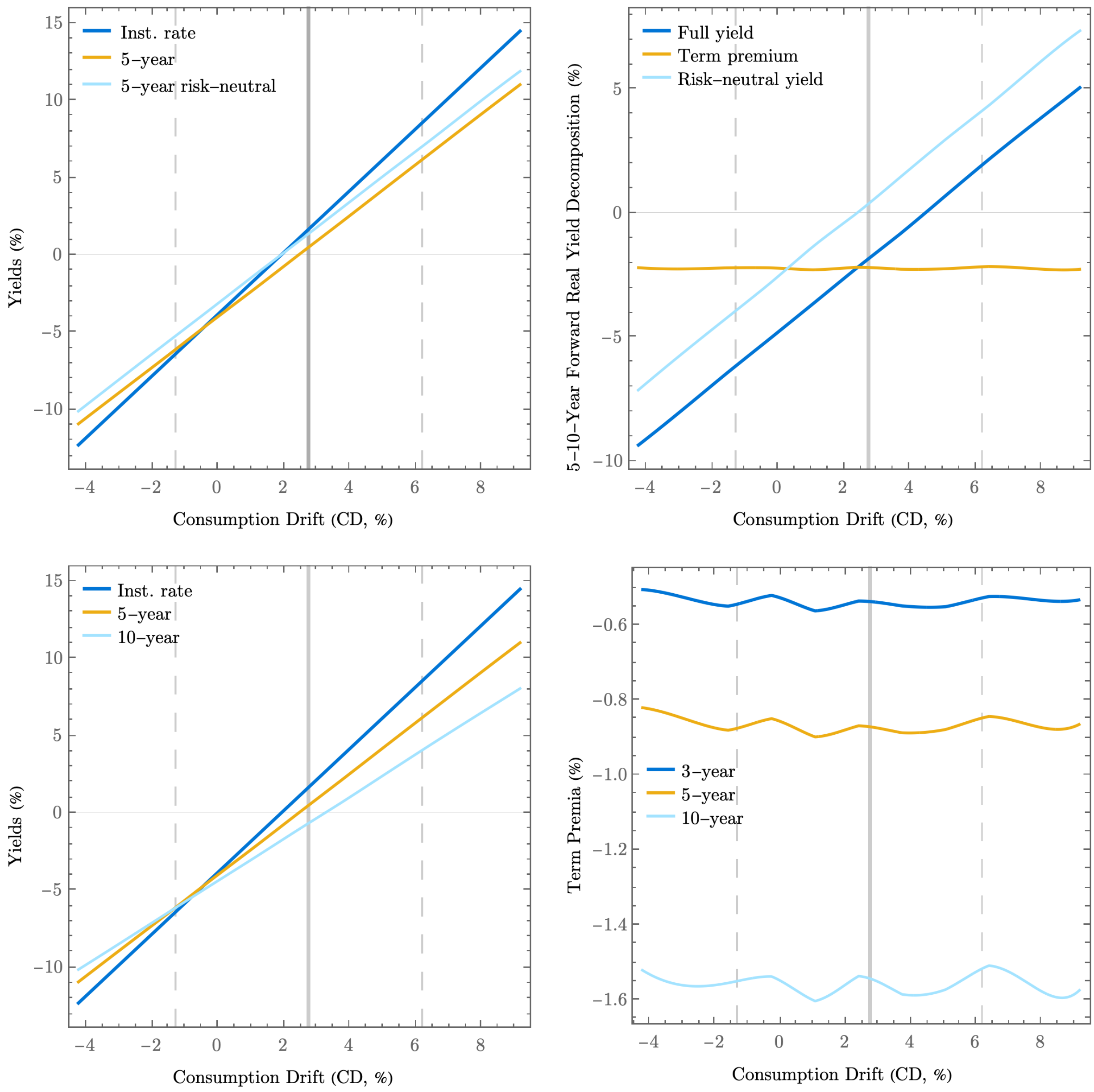
range$_{\text{low}}$(r$_{2,\text{model}}$)=0.0977548, range$_{\text{high}}$(r$_{2,\text{model}}$)=0.183359, range(r$_{2,\text{data}}$)=dataRange.

Figure 21: Time-varying and high CD with TSU.
See Figure 16 for more details about the plots.

(variation overview)

F.8 TSU-CD-HCV, $\sigma_{ct} = 0.16$

Yields move down and term premia increase in absolute value, but they are again constant.



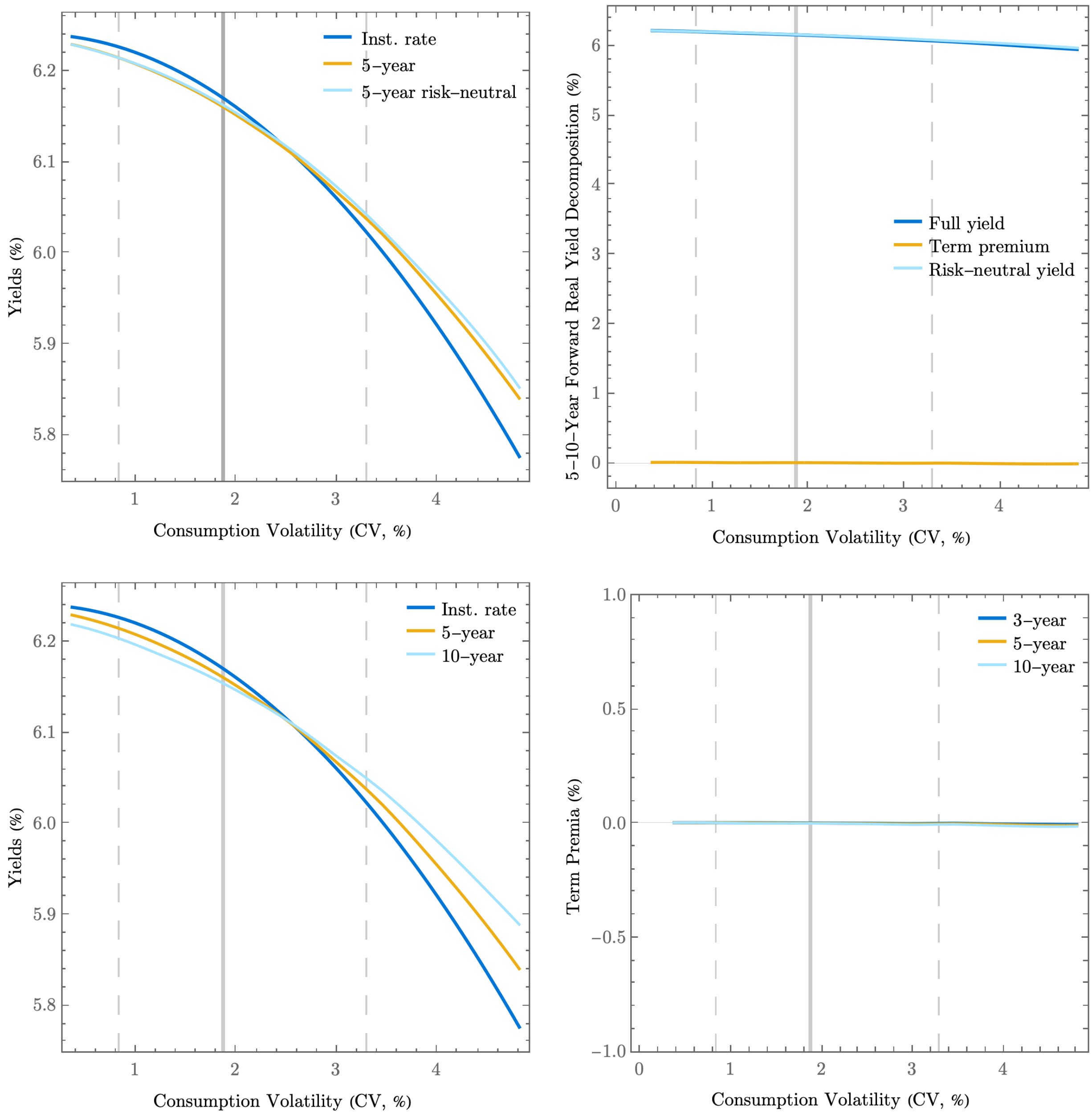
γ=2, ρ=0.012, μ_{ct}=x_t + 0.0252, σ_{ct}=0.16, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5

range$_{\mathrm{low}}$(r$_{2,\mathrm{model}}$)=0.107642, range$_{\mathrm{high}}$(r$_{2,\mathrm{model}}$)=0.207586, range(r$_{2,\mathrm{data}}$)=0.071748.

Figure 22: Time-varying CD with TSU and HCV.
See Figure 16 for more details about the plots.

(variation overview)

F.9 TSU-CV, Calibration used in main paper, Figure 3



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02 x_t, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.21 $\sqrt{x_t}$, ρ_{cx}=−0.5

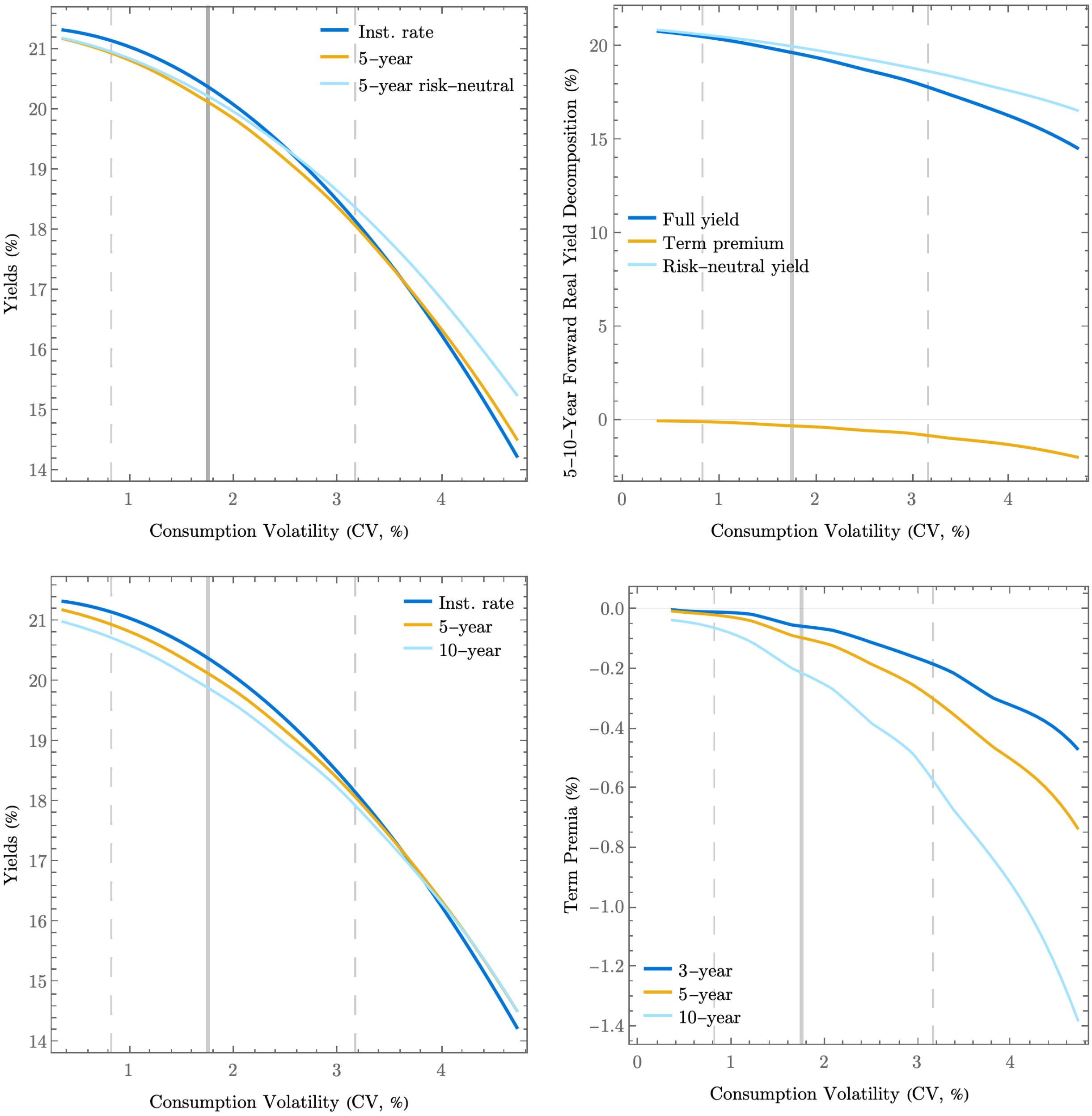
range$_{\text{low}}$(r$_{2,\text{model}}$)=0.00103855, range$_{\text{high}}$(r$_{2,\text{model}}$)=0.00408583, range(r$_{2,\text{data}}$)=0.071748.

Figure 23: Time-varying CV with TSU.
See Figure 16 for more details about the plots.

(variation overview)

F.10 TSU-CV-HRA, $\gamma = 8$

Term premia increased in absolute value but not enough and yields moved very high.



γ=8, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02 x_t, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.21 $\sqrt{x_t}$, ρ_{cx}=−0.5

range$_{low}$(r$_{2,model}$)=0.0157495, range$_{high}$(r$_{2,model}$)=0.0684834, range(r$_{2,data}$)=0.071748.

Figure 24: Time-varying CV with high risk aversion.
See Figure 16 for more details about the plots.

(variation overview)

F.11 TSU-CV-HCD, $\mu_c 0 = 0.08$

Term premia did not change but yields move implausibly high.

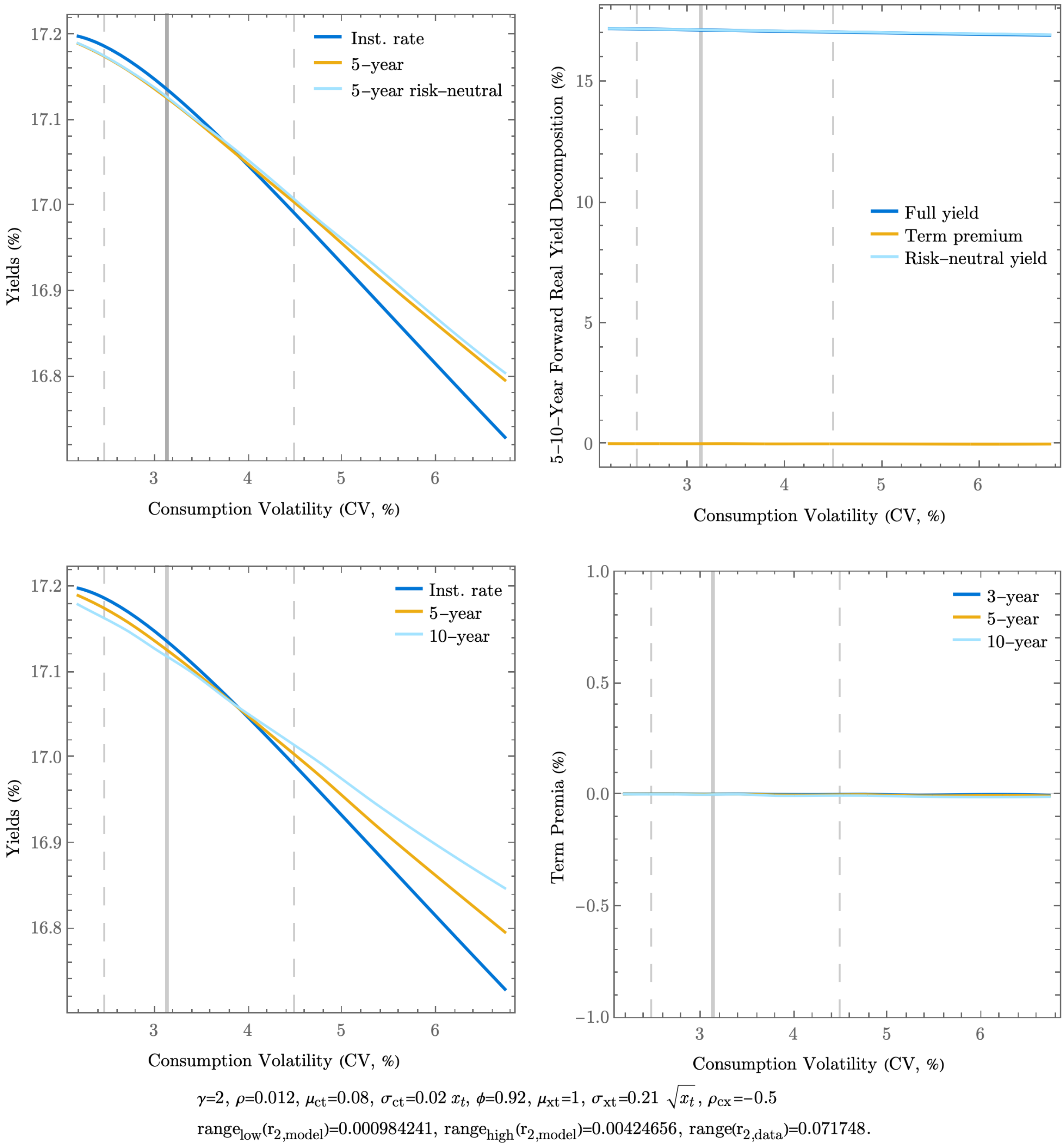
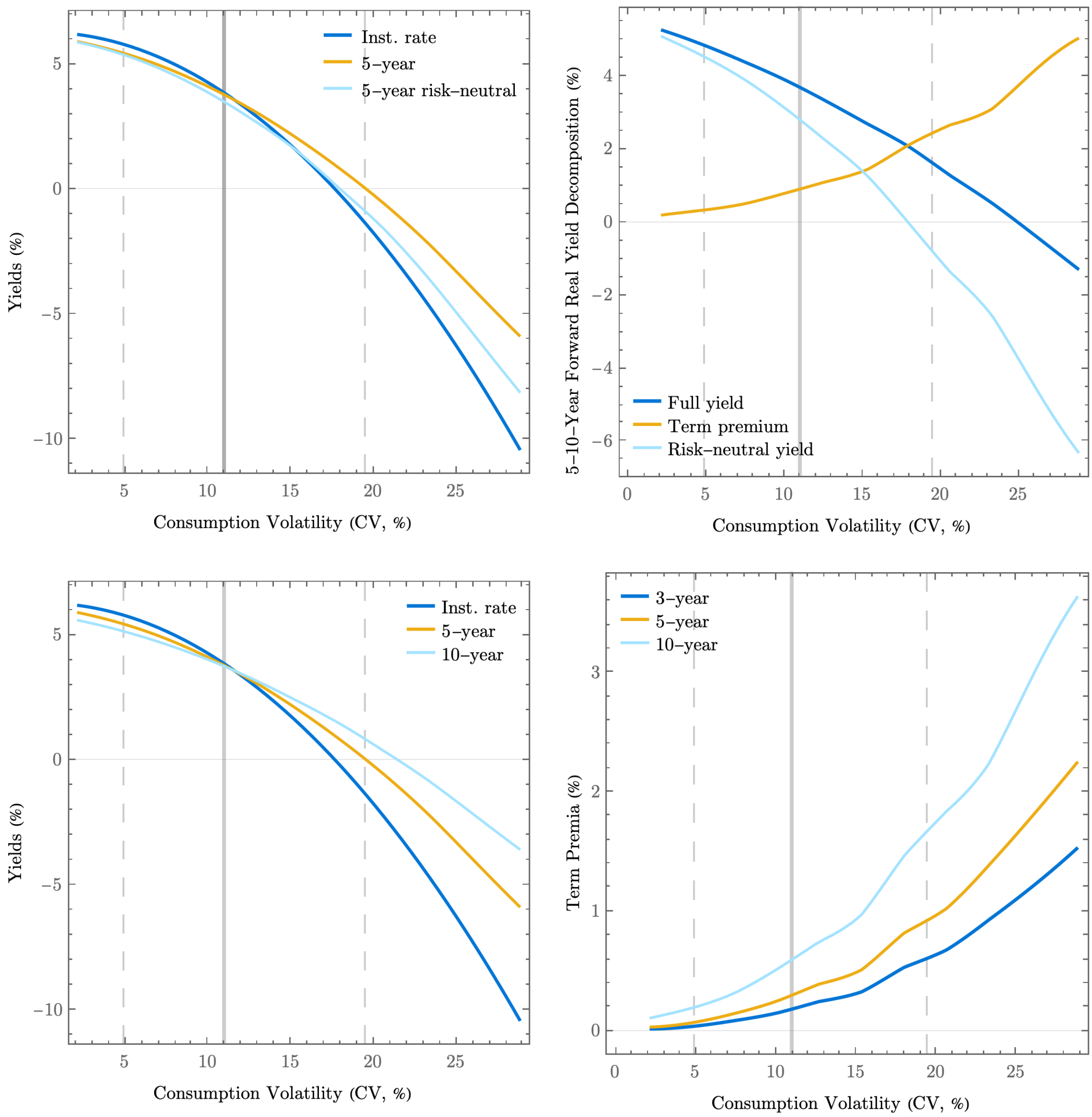


Figure 25: Time-varying CV with TSU and high CD.
See Figure 16 for more details about the plots.

(variation overview)

F.12 TSU-HCV, Calibration used in main paper, Figure 6



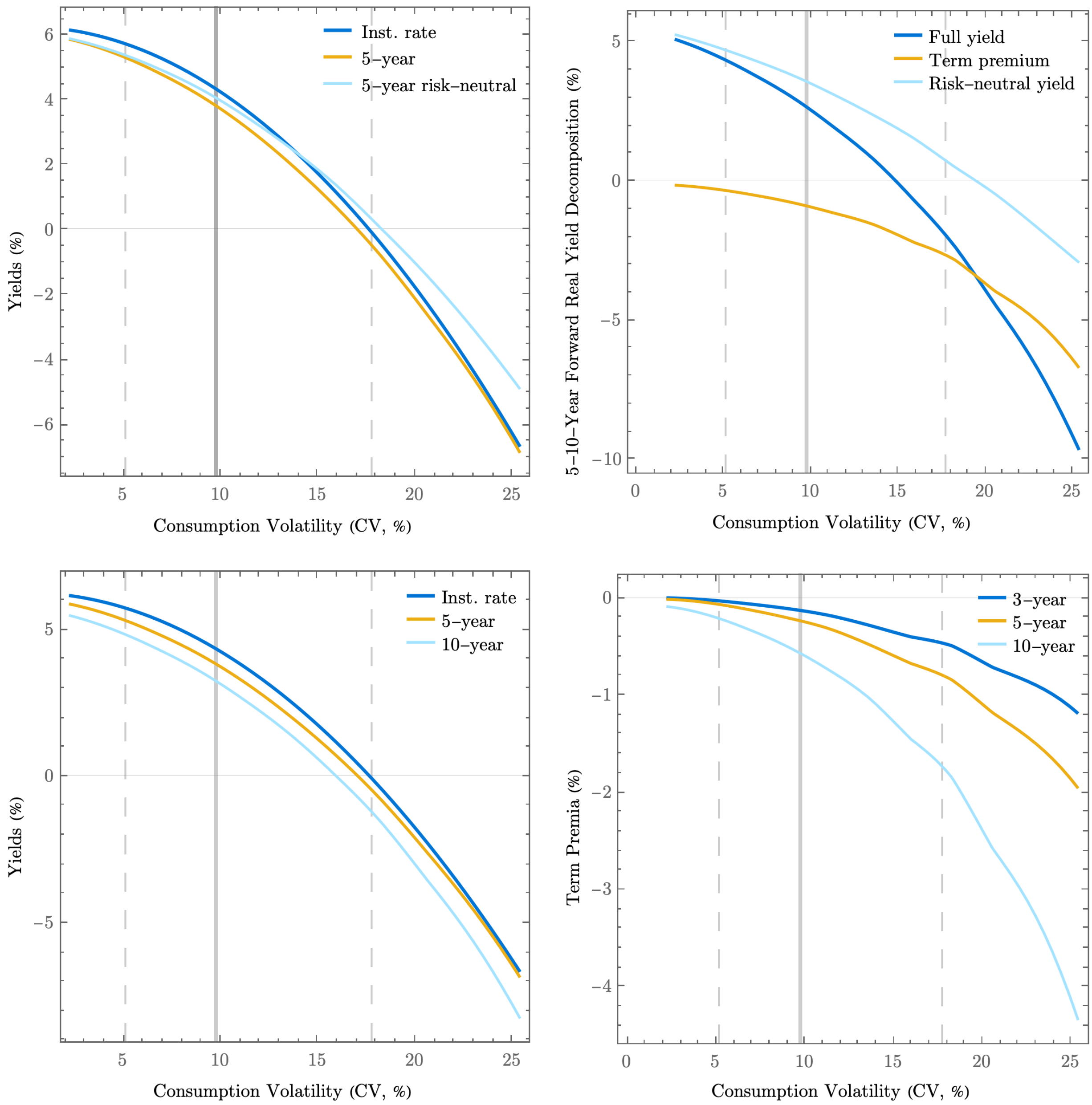
γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.12 x_t, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.21 $\sqrt{x_t}$, ρ_{cx}=0.5

$\text{range}_{\text{low}}(\text{r}_{2,\text{model}})$=0.0374496, $\text{range}_{\text{high}}(\text{r}_{2,\text{model}})$=0.149538, $\text{range}(\text{r}_{2,\text{data}})$=0.071748.

Figure 26: Time-varying and HCV with TSU and positive ρ_{cx}.
See Figure 16 for more details about the plots.

(variation overview)

F.13 TSU-HCV-NCor, $\rho_{cx} < 0$



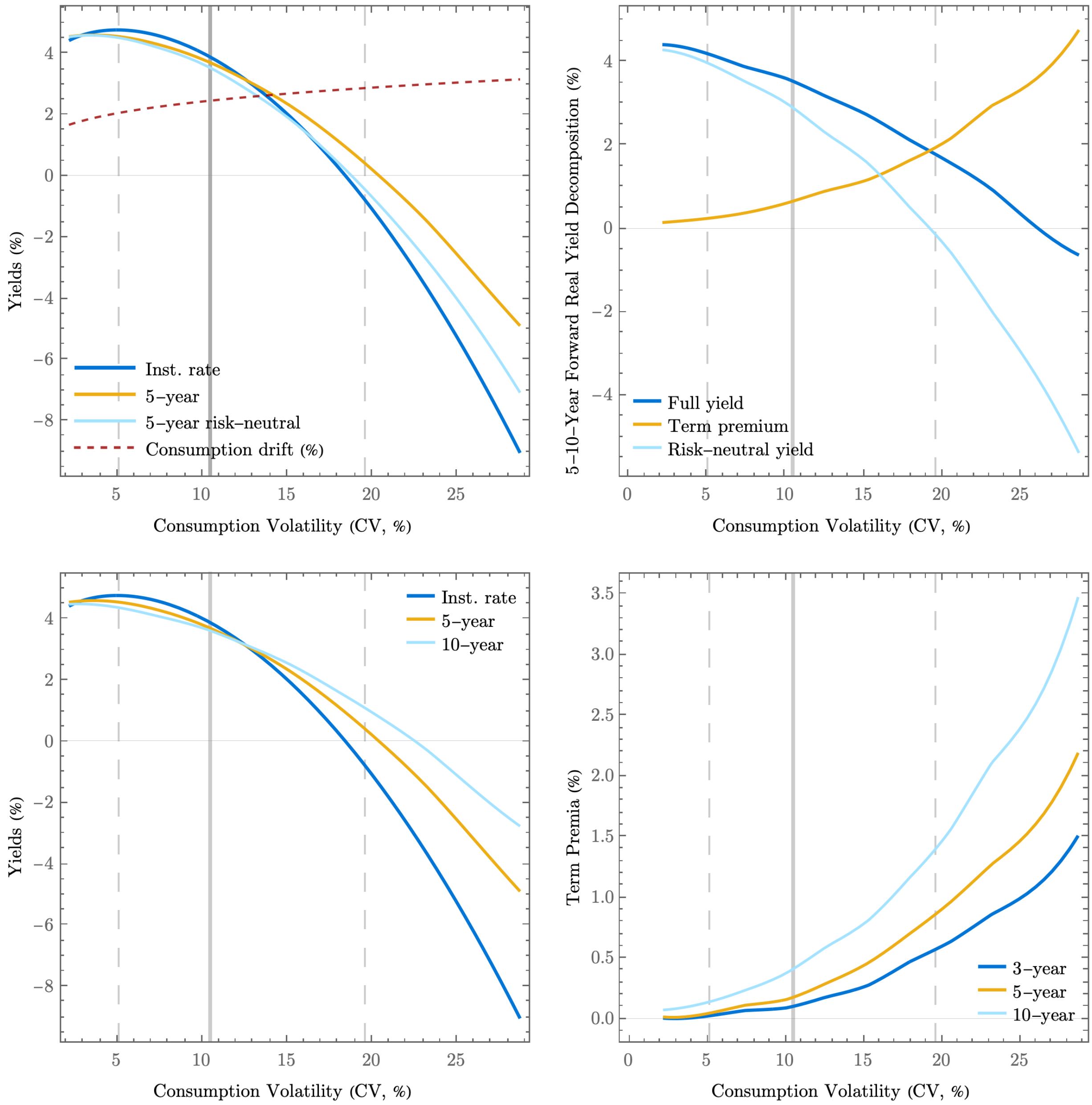
γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.11 x_t, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.2 $\sqrt{x_t}$, ρ_{cx}=−0.5

range$_{\text{low}}$(r$_{\text{2,model}}$)=0.0306062, range$_{\text{high}}$(r$_{\text{2,model}}$)=0.119326, range(r$_{\text{2,data}}$)=0.071748.

Figure 27: Time-varying CV with TSU and negative ρ_{cx}.
See Figure 16 for more details about the plots.

(variation overview)

F.14 Arb-DP, Calibration used in main paper, Figure 6



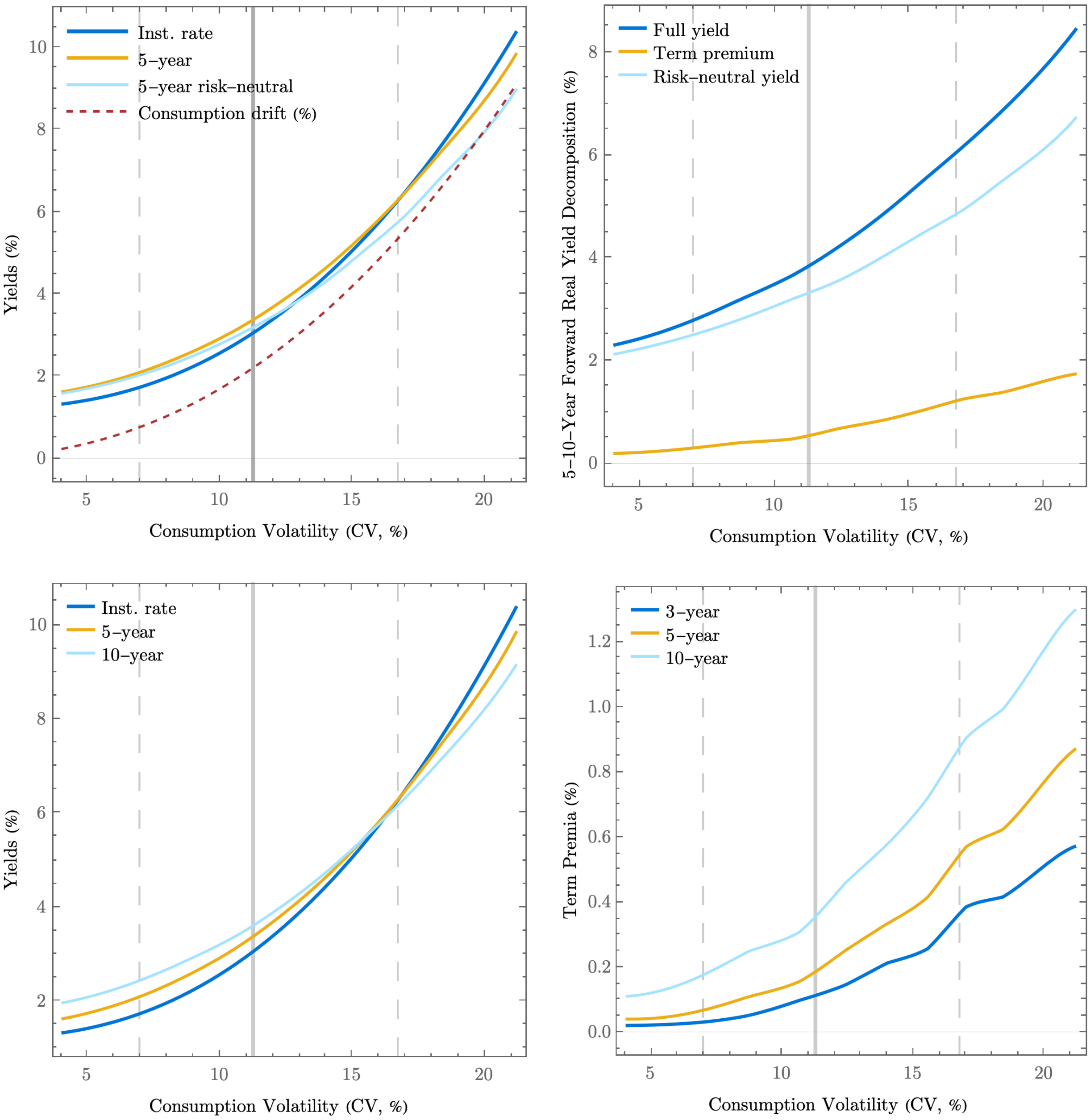
γ=2, ρ=0.012, $\mu_{\rm ct}$=0.0252 $\sqrt[4]{x_t}$, $\sigma_{\rm ct}$=0.12 x_t, ϕ=0.92, $\mu_{\rm xt}$=1, $\sigma_{\rm xt}$=0.21 $\sqrt{x_t}$, $\rho_{\rm cx}$=0.5

range$_{\rm low}$(r$_{\rm 2,model}$)=0.0222875, range$_{\rm high}$(r$_{\rm 2,model}$)=0.132065, range(r$_{\rm 2,data}$)=0.071748.

Figure 28: Both time-varying CD and CV with TSU, short-term rate decreasing in CV and positive ρ_{cx}.
See Figure 6 for more details about the plots.

(variation overview)

F.15 Arb-IN, Calibration used in main paper, Figure 6



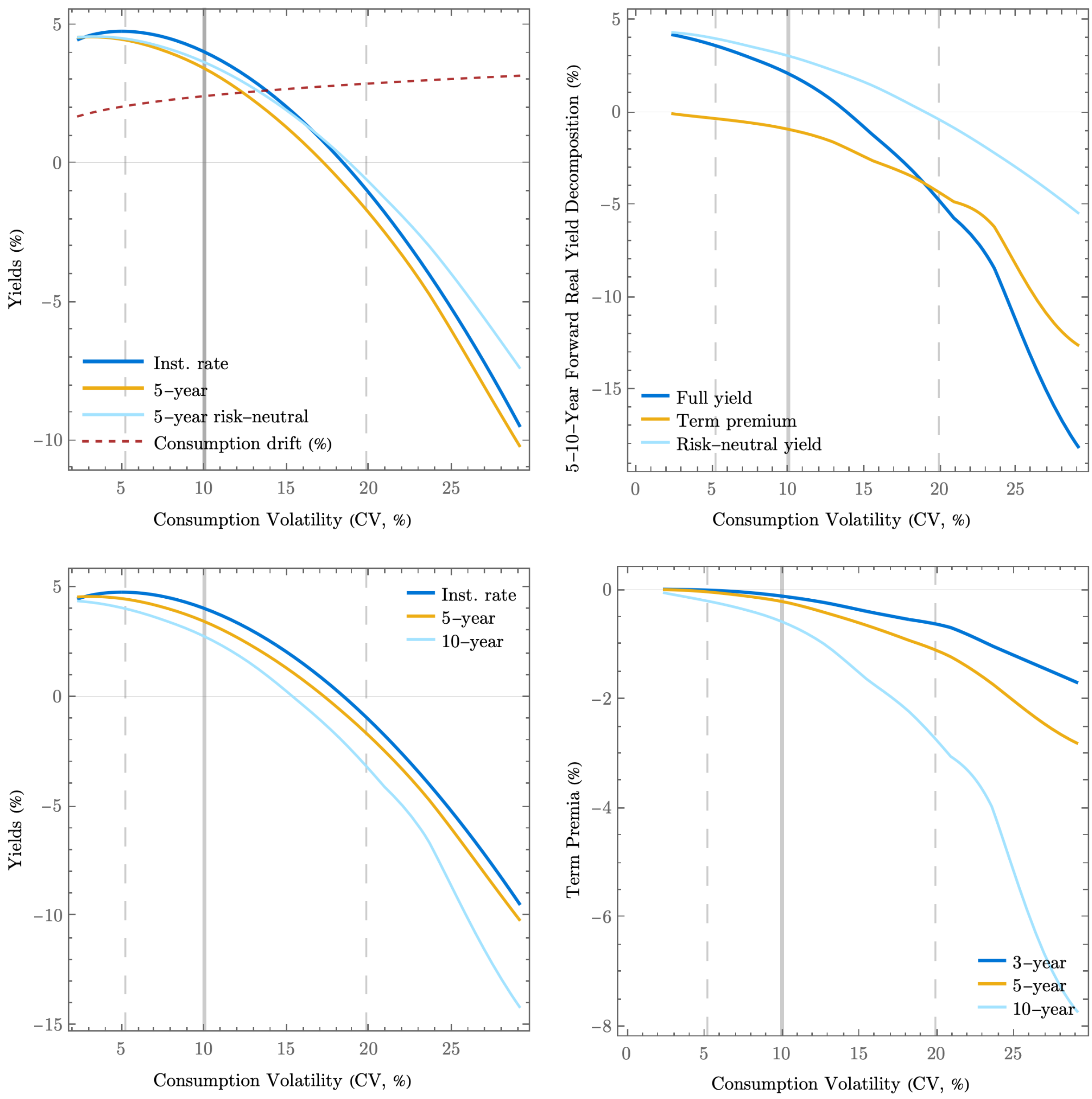
γ=2, ρ=0.012, μ_{ct}=0.0252 $x_t^{3/2}$, σ_{ct}=0.12 $x_t^{2/3}$, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.21 $\sqrt{x_t}$, ρ_{cx}=−0.5

$\text{range}_{\text{low}}(\text{r}_{2,\text{model}})$=0.0258998, $\text{range}_{\text{high}}(\text{r}_{2,\text{model}})$=0.0818517, $\text{range}(\text{r}_{2,\text{data}})$=dataRange.

Figure 29: Both time-varying CD and CV with TSU, short-term rate increasing in CV and negative ρ_{cx}.
See Figure 6 for more details about the plots.

(variation overview)

F.16 Arb-DN



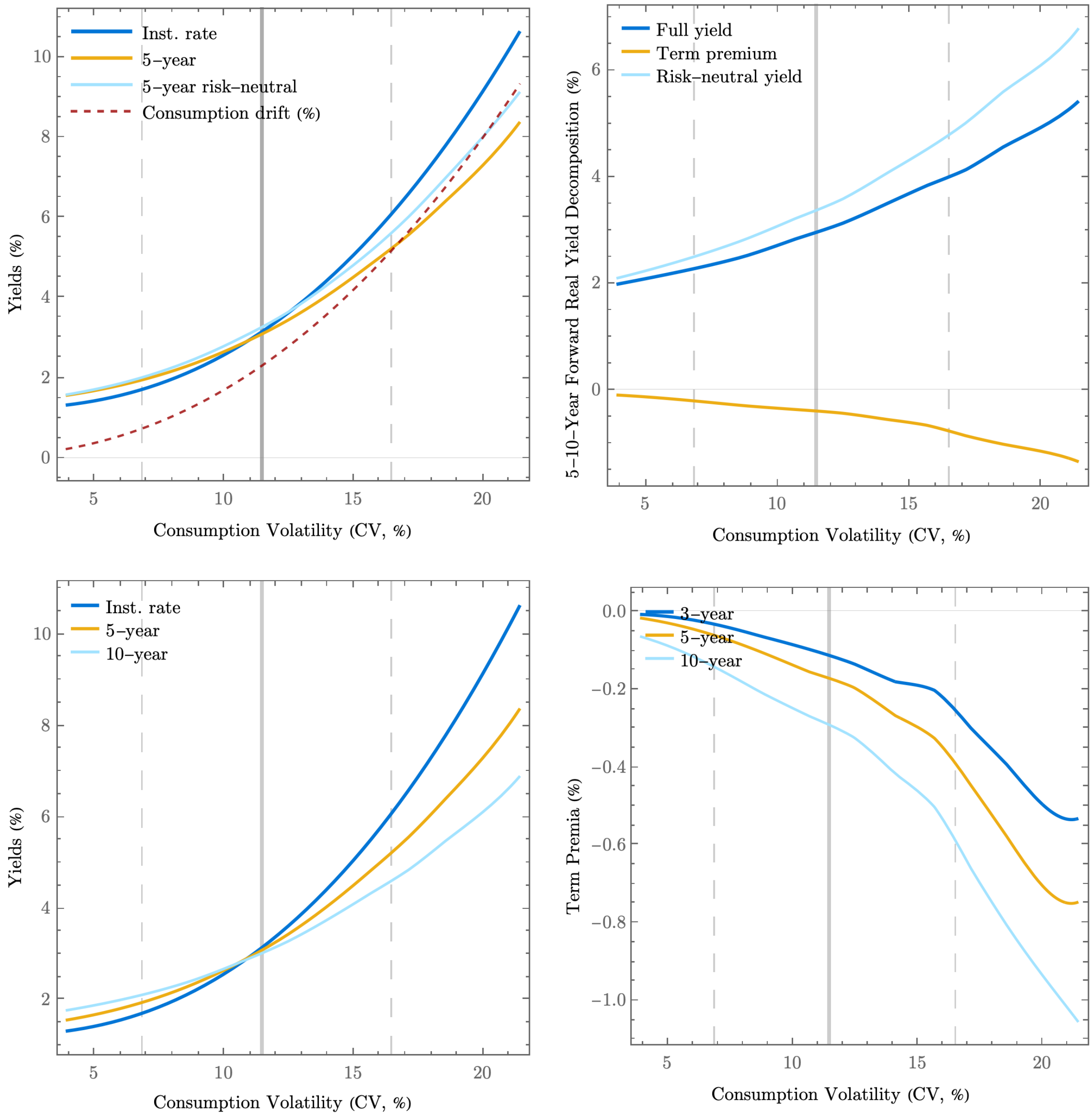
γ=2, ρ=0.012, μ_{ct}=0.0252 $\sqrt[4]{x_t}$, σ_{ct}=0.12 x_t, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.21 $\sqrt{x_t}$, ρ_{cx}=−0.5

$\mathrm{range_{low}}(\mathrm{r_{2,model}})$=0.0224998, $\mathrm{range_{high}}(\mathrm{r_{2,model}})$=0.134553, $\mathrm{range}(\mathrm{r_{2,data}})$=0.071748.

Figure 30: Both time-varying CD and CV with TSU, short-term rate decreasing in CV and negative ρ_{cx}.

(variation overview)

F.17 Arb-IP



γ=2, ρ=0.012, μ_{ct}=0.0252 $x_t^{3/2}$, σ_{ct}=0.12 $x_t^{2/3}$, ϕ=0.92, μ_{xt}=1, σ_{xt}=0.21 $\sqrt{x_t}$, ρ_{cx}=0.5

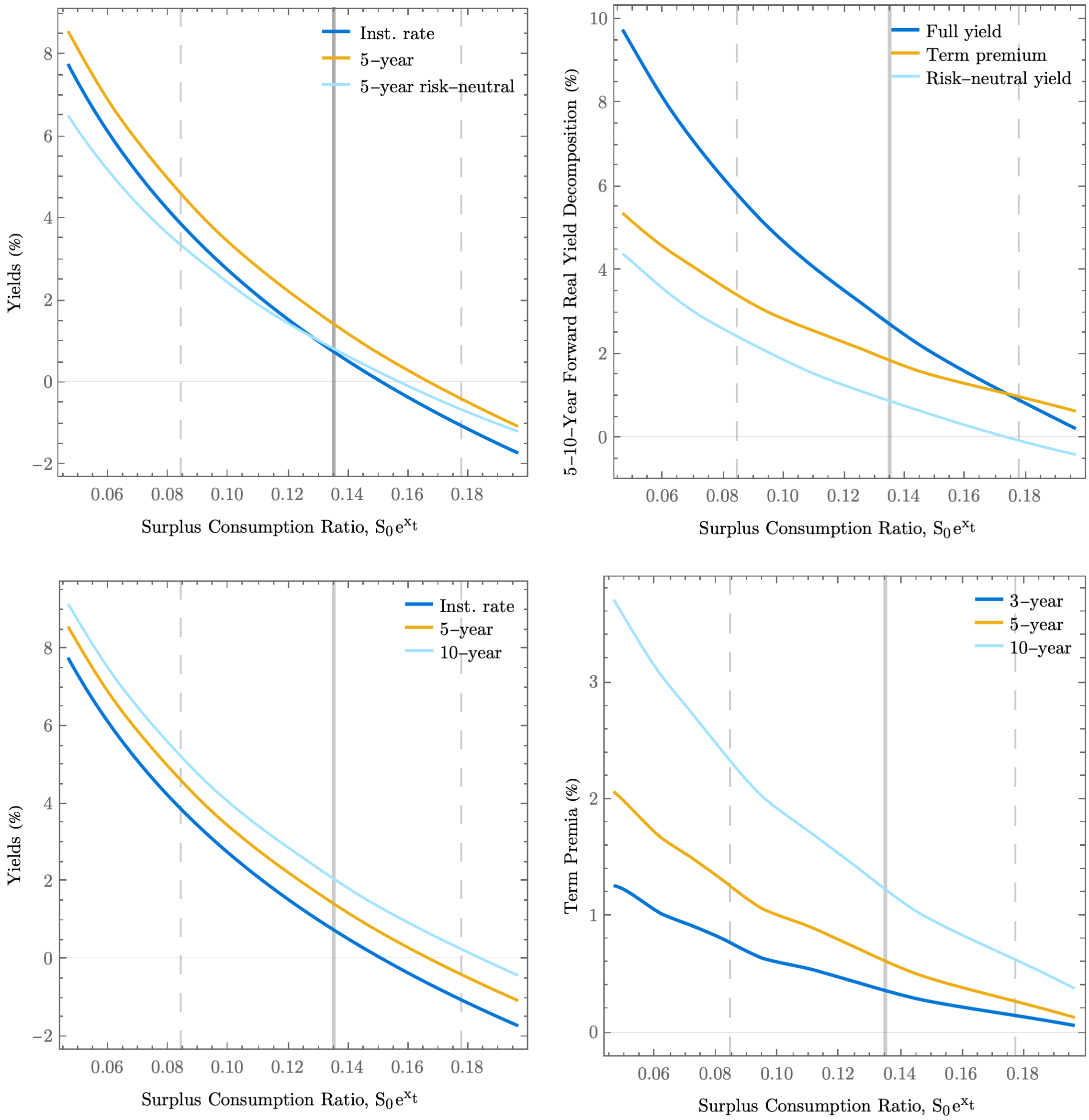
$\mathrm{range_{low}(r_{2,model})}$=0.0263095, $\mathrm{range_{high}(r_{2,model})}$=0.0794248, $\mathrm{range(r_{2,data})}$=0.071748.

Figure 31: Both time-varying CD and CV with TSU, short-term rate increasing in CV and positive ρ_{cx}.
See Figure 6 for more details about the plots.

(variation overview)

F.18 TSU-Habit

F.19 Calibration used in main paper, Figure 3



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.02 $\mathrm{If}\left[x_t > 0.492061,\, 0,\, \dfrac{\sqrt{1-2\,x_t}}{0.02\sqrt{\dfrac{2}{-\frac{0.066}{2}-\log(0.92)}}} - 1\right]$, ρ_{cx}=1

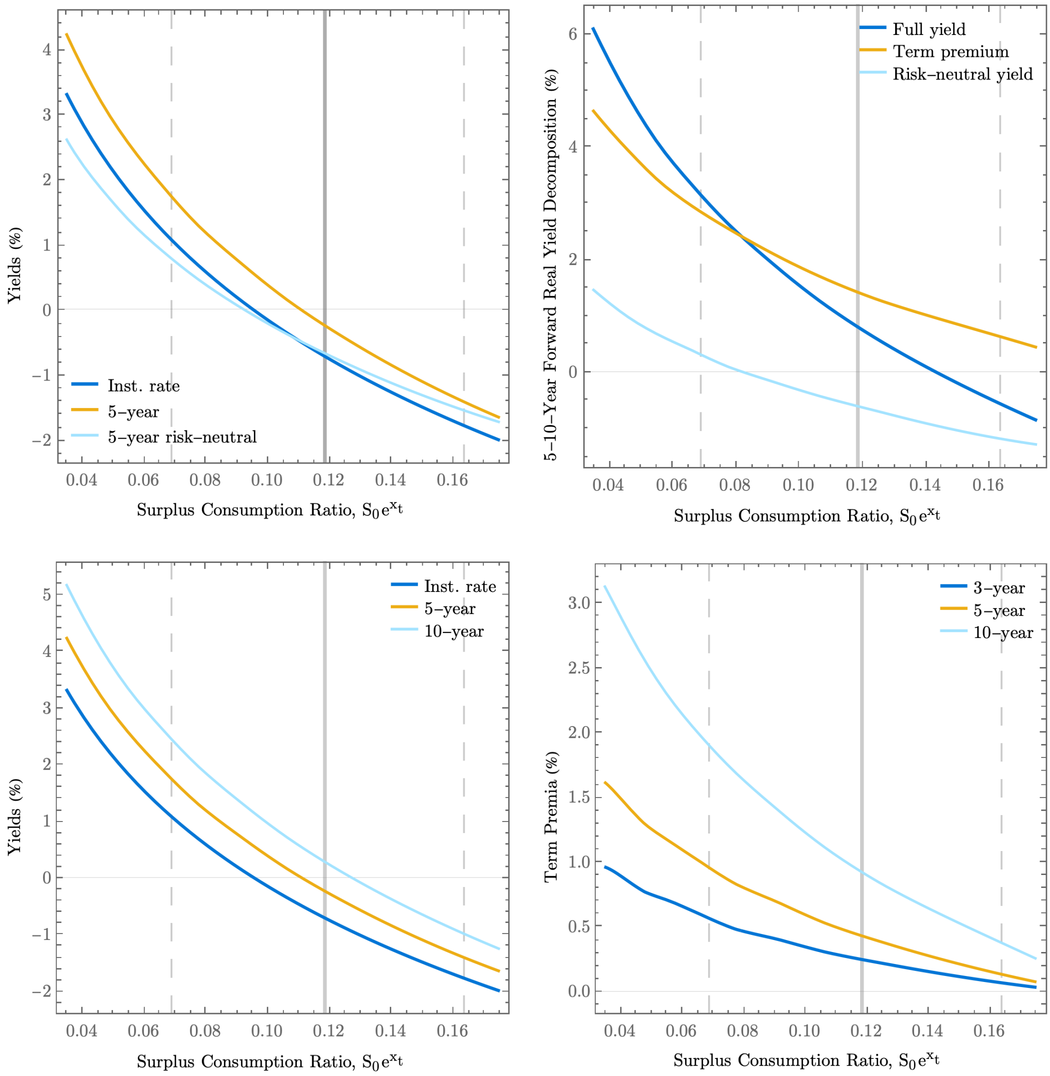
$\mathrm{range}_{\mathrm{low}}(\mathrm{r}_{2,\mathrm{model}})$=0.0323706, $\mathrm{range}_{\mathrm{high}}(\mathrm{r}_{2,\mathrm{model}})$=0.0831215, $\mathrm{range}(\mathrm{r}_{2,\mathrm{data}})$=dataRange.

Figure 32: Time-varying external habit with TSU.
See Figure 16 for more details about the plots.

(variation overview)

F.20 TSU-Habit-Low.b, $b = 0.033$

Term premia did not change but yields became flatter. This is noteworthy because in Abrahams *et al.* (2016) forward term premia are big while the forward risk-neutral yields are small in absolute value.



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, $\sigma_{xt}=0.02\,\mathrm{If}\left[x_t > 0.494019,\ 0,\ \frac{\sqrt{1-2\,x_t}}{0.02\sqrt{\frac{2}{-\frac{0.033}{2}-\log(0.92)}}}-1\right]$, ρ_{cx}=1

$\mathrm{range}_{\mathrm{low}}(r_{2,\mathrm{model}})=0.0181377$, $\mathrm{range}_{\mathrm{high}}(r_{2,\mathrm{model}})=0.0475884$, $\mathrm{range}(r_{2,\mathrm{data}})=\mathrm{dataRange}$.

Figure 33: Time-varying external habit with TSU.
See Figure 16 for more details about the plots.

(variation overview)

F.21 TSU-Habit-Neg.b, $b = -0.033$

The short-term rate is now pro-cyclical and term premia are negative.

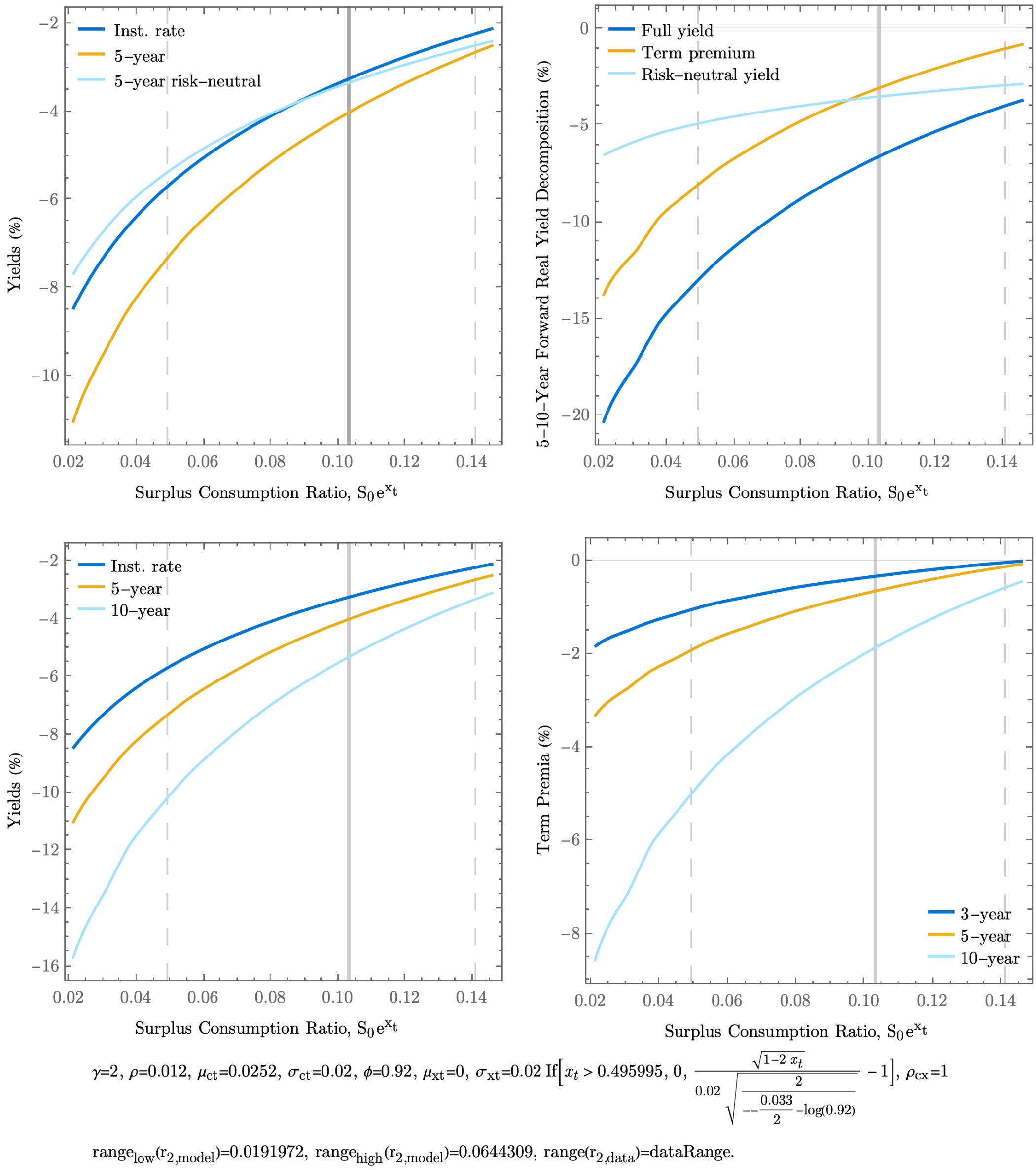
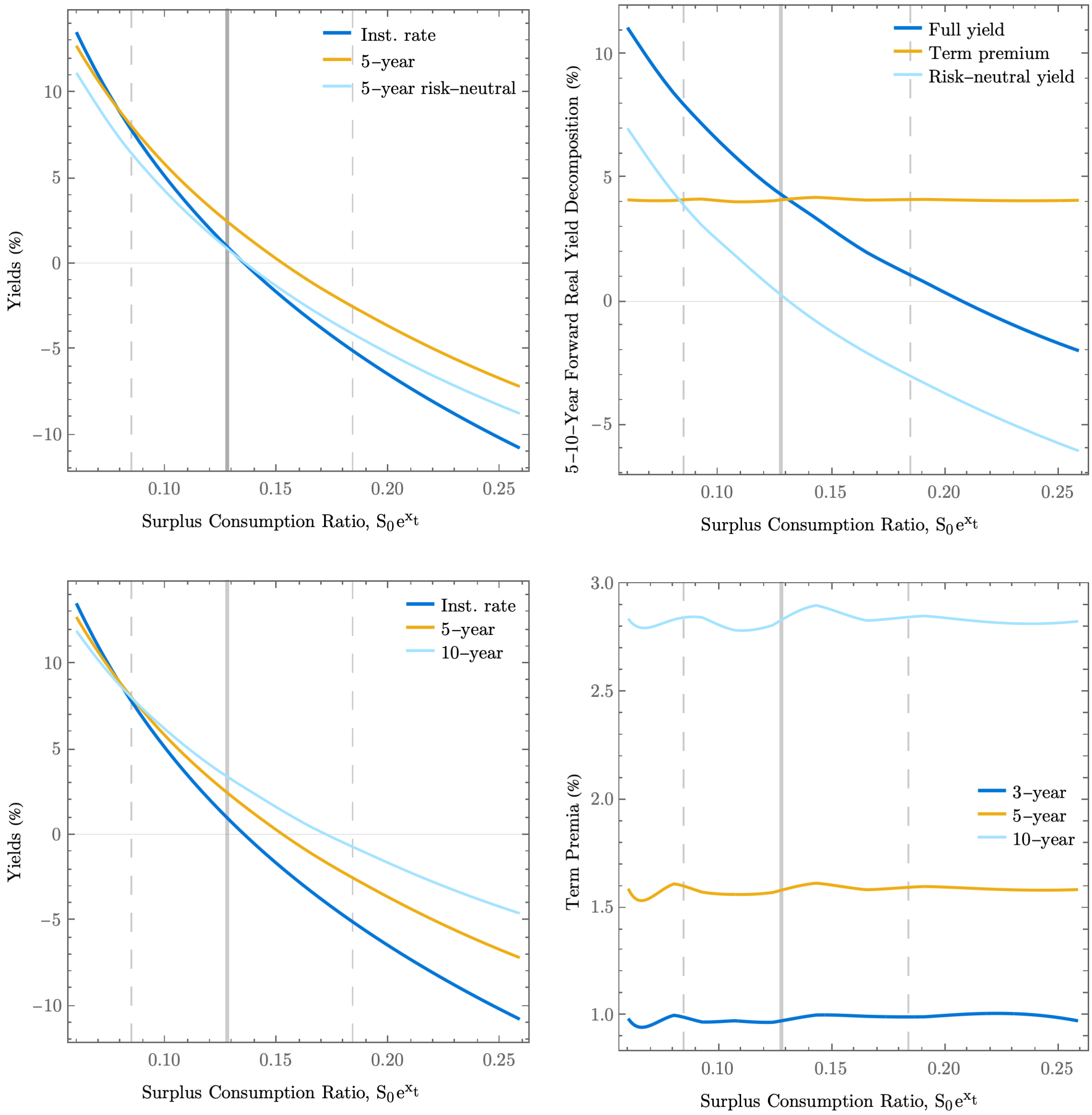


Figure 34: Time-varying external habit with TSU and $b < 0$.
See Figure 16 for more details about the plots.

(variation overview)

F.22 TSU-Habit-CSV, $\sigma_{xt} = \lambda(0)\sigma_{c0}$

The term premia are now constant. This is partially contrary to the spirit of Campbell and Cochrane (1999), because the surplus consumption ratio does not get more volatile in bad states of the economy, but it illustrates how heteroskedasticity is crucial for the generation of variable term premia.



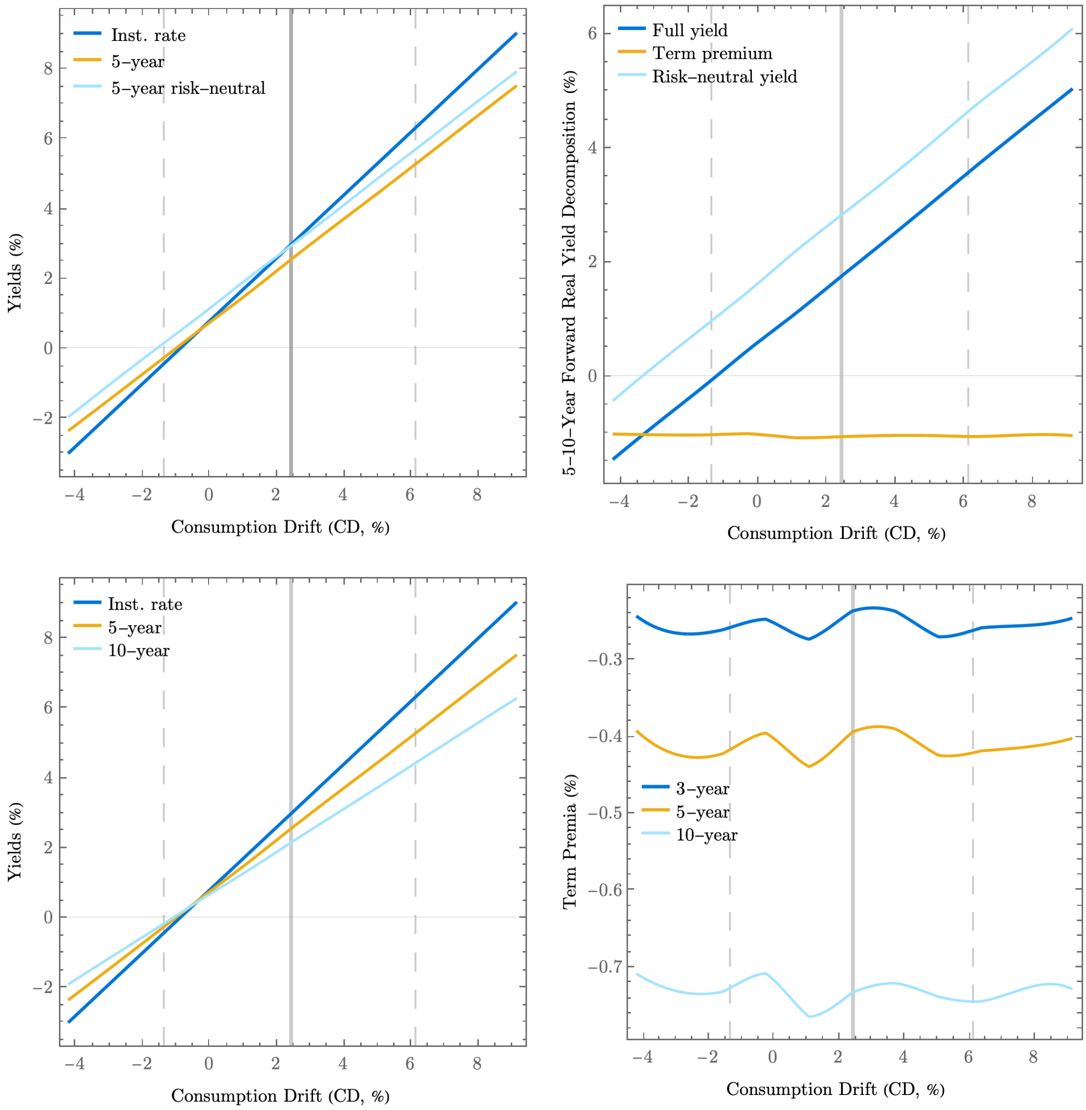
γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.138716, ρ_{cx}=1

$\text{range}_{\text{low}}(r_{2,\text{model}})$=0.103559, $\text{range}_{\text{high}}(r_{2,\text{model}})$=0.185687, $\text{range}(r_{2,\text{data}})$=dataRange.

Figure 35: Time-varying external habit with TSU and constant state variable volatility.
See Figure 16 for more details about the plots.

(variation overview)

F.23 RU-CD, Calibration used in main paper, Figure 5



γ=2, ρ=0.012, $\mu_{\mathrm{ct}}=x_t+0.0252$, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5, ψ=1.11, approx. order=8

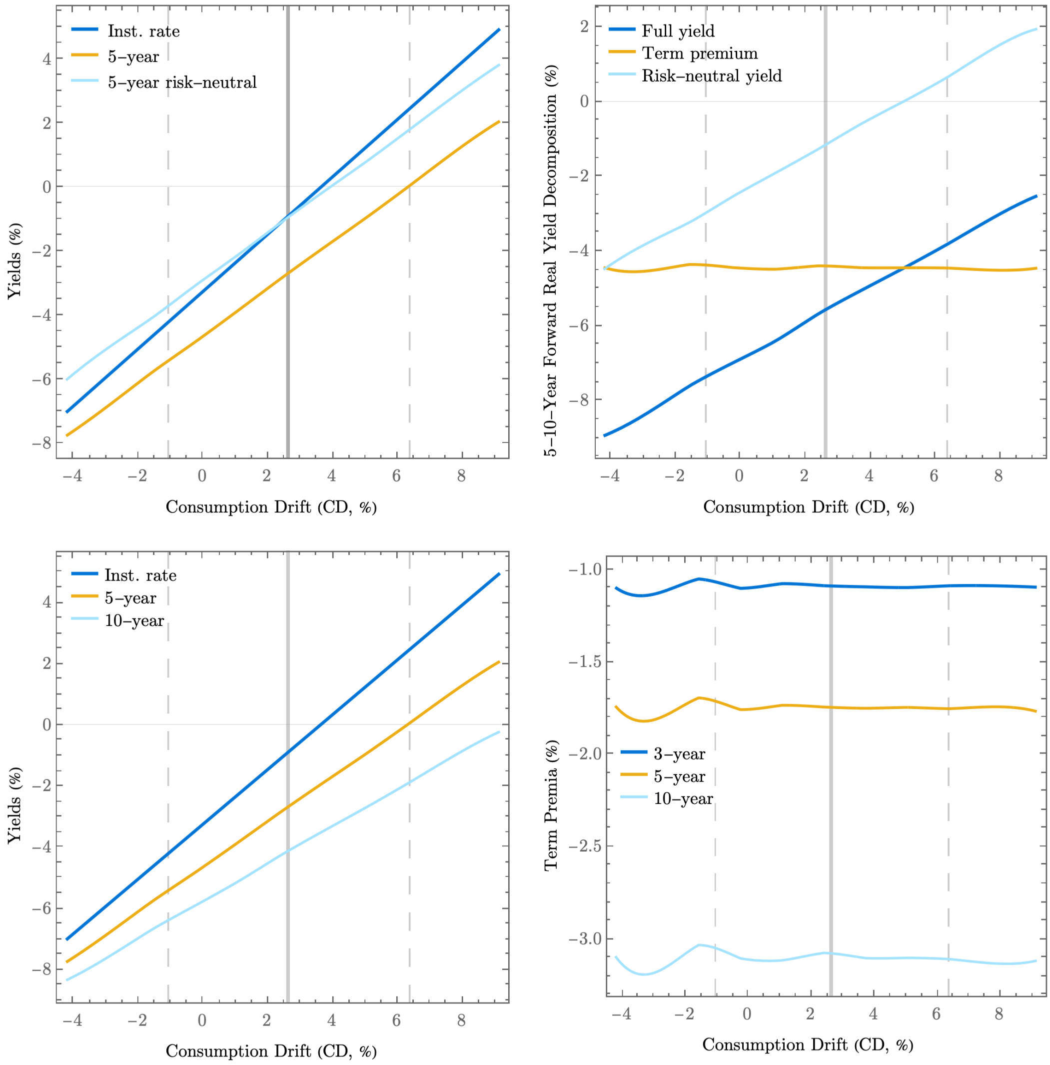
$\mathrm{range_{low}}(\mathrm{r_{2,model}})$=0.050672, $\mathrm{range_{high}}(\mathrm{r_{2,model}})$=0.0927875, $\mathrm{range}(\mathrm{r_{2,data}})$=dataRange.

Figure 36: Time-varying CD with RU.
See Figure 16 for more details about the plots.

(variation overview)

F.24 RU-CD-HRA, $\gamma = 6$

Term premia stay constant and negative but they become significantly larger in absolute value. In this paper a time-varying γ parameter is not considered, but this suggests that a time-varying risk aversion would be able to produce time-varying term premia. The habit model essentially provides a similar mechanism.



γ=6, ρ=0.012, μ_{ct}=x_t+0.0252, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5, ψ=1.11, approx. order=8

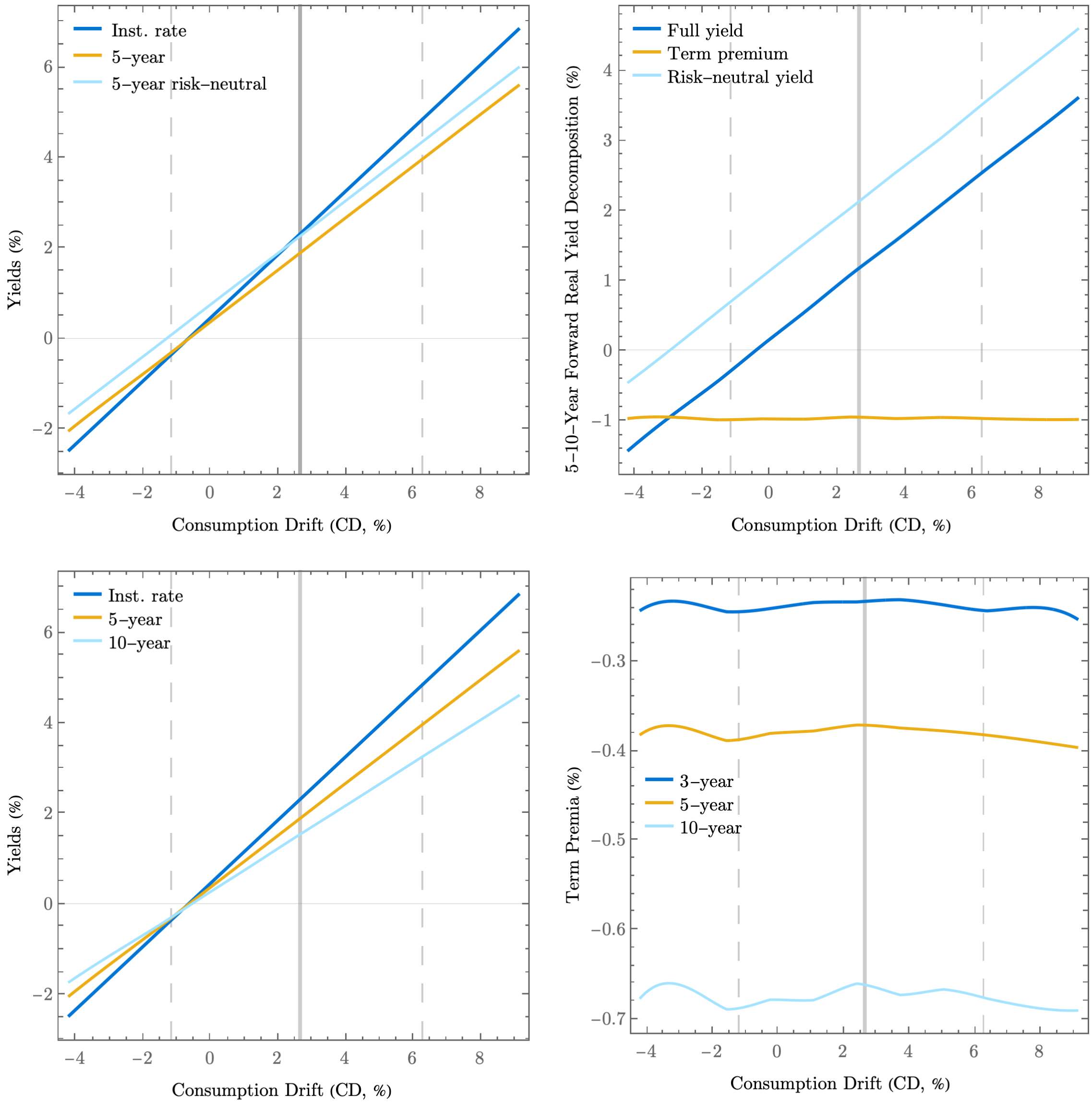
$\text{range}_{\text{low}}(\text{r}_{2,\text{model}})$=0.0491264, $\text{range}_{\text{high}}(\text{r}_{2,\text{model}})$=0.0893581, $\text{range}(\text{r}_{2,\text{data}})$=0.071748.

Figure 37: Time-varying CD with RU and high risk aversion.
See Figure 5 for more details about the plots.

(variation overview)

F.25 RU-CD-HIES, $\psi = 1.43$

Term premia do not seem to change significantly. The range of the short rate increases.



γ=2, ρ=0.012, $\mu_{\mathrm{ct}}=x_t+0.0252$, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5, ψ=1.43, approx. order=8

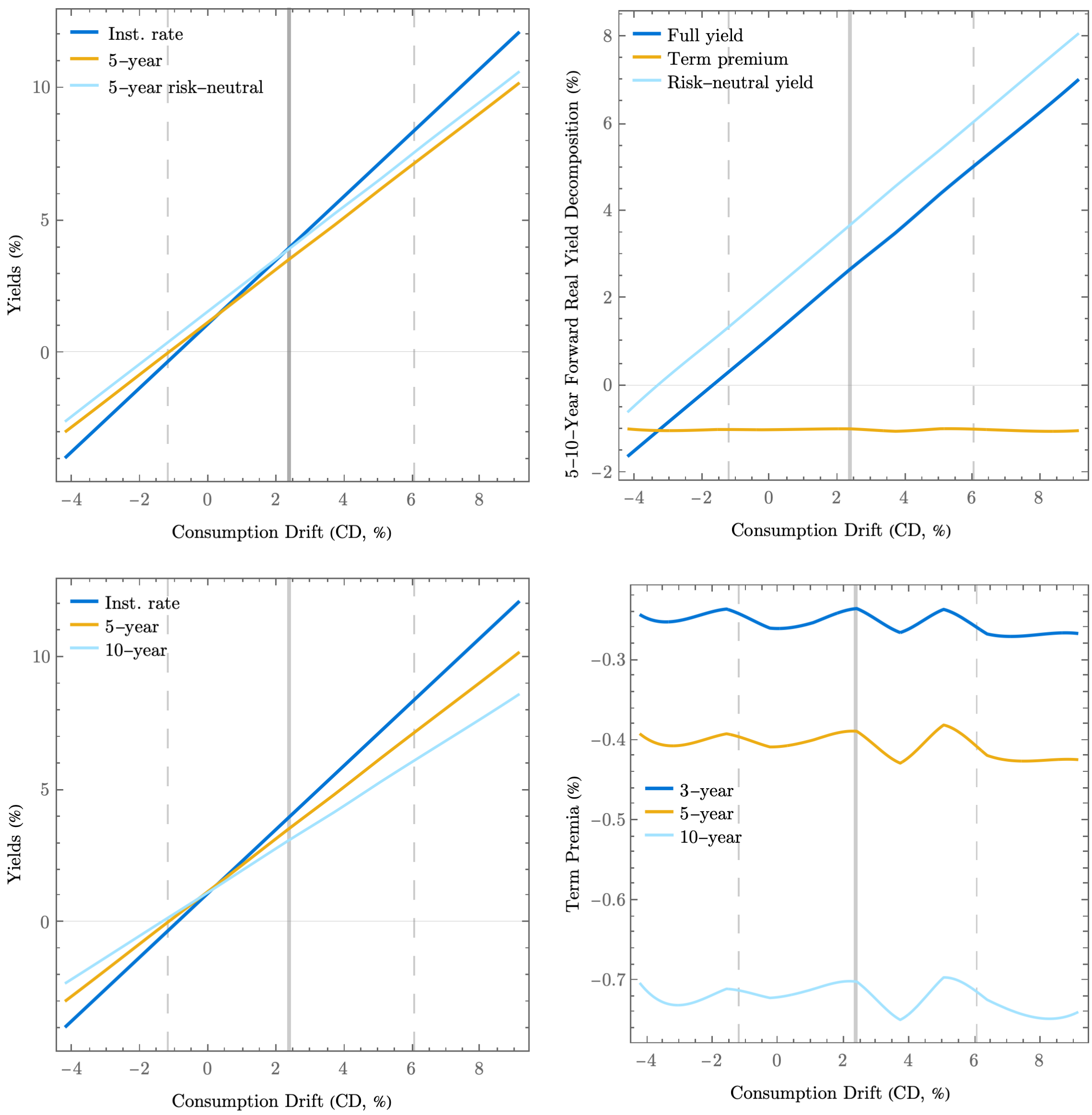
$\mathrm{range_{low}}(\mathrm{r_{2,model}})$=0.037872, $\mathrm{range_{high}}(\mathrm{r_{2,model}})$=0.071687, $\mathrm{range}(\mathrm{r_{2,data}})$=0.071748.

Figure 38: Time-varying CD with HIES.
See Figure 16 for more details about the plots.

(variation overview)

F.26 RU-CD-LIES, $\psi = 0.83$

Term premia do not seem to change significantly. Curiously the range of the short rate increases again.



γ=2, ρ=0.012, $\mu_{ct}=x_t+0.0252$, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5, ψ=0.83, approx. order=11

$\text{range}_{\text{low}}(r_{2,\text{model}})$=0.0651376, $\text{range}_{\text{high}}(r_{2,\text{model}})$=0.121441, $\text{range}(r_{2,\text{data}})$=0.071748.

Figure 39: Time-varying CD with RU with LIES.
See Figure 16 for more details about the plots.

(variation overview)

F.27 RU-CD-HCor, $\rho_{cx} = 1$

Term premia increase in absolute value but do not double in size as did the correlation parameter.

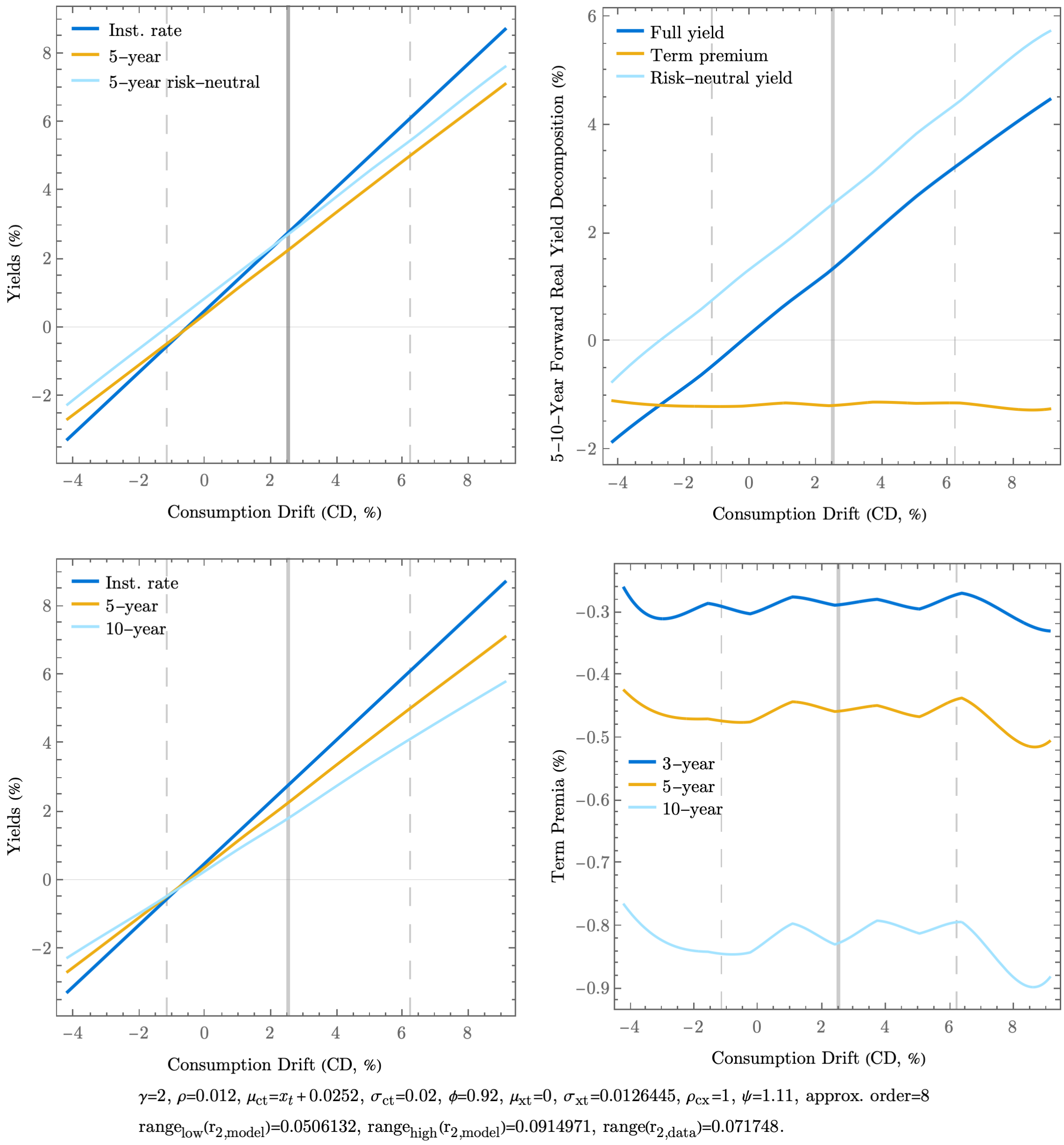
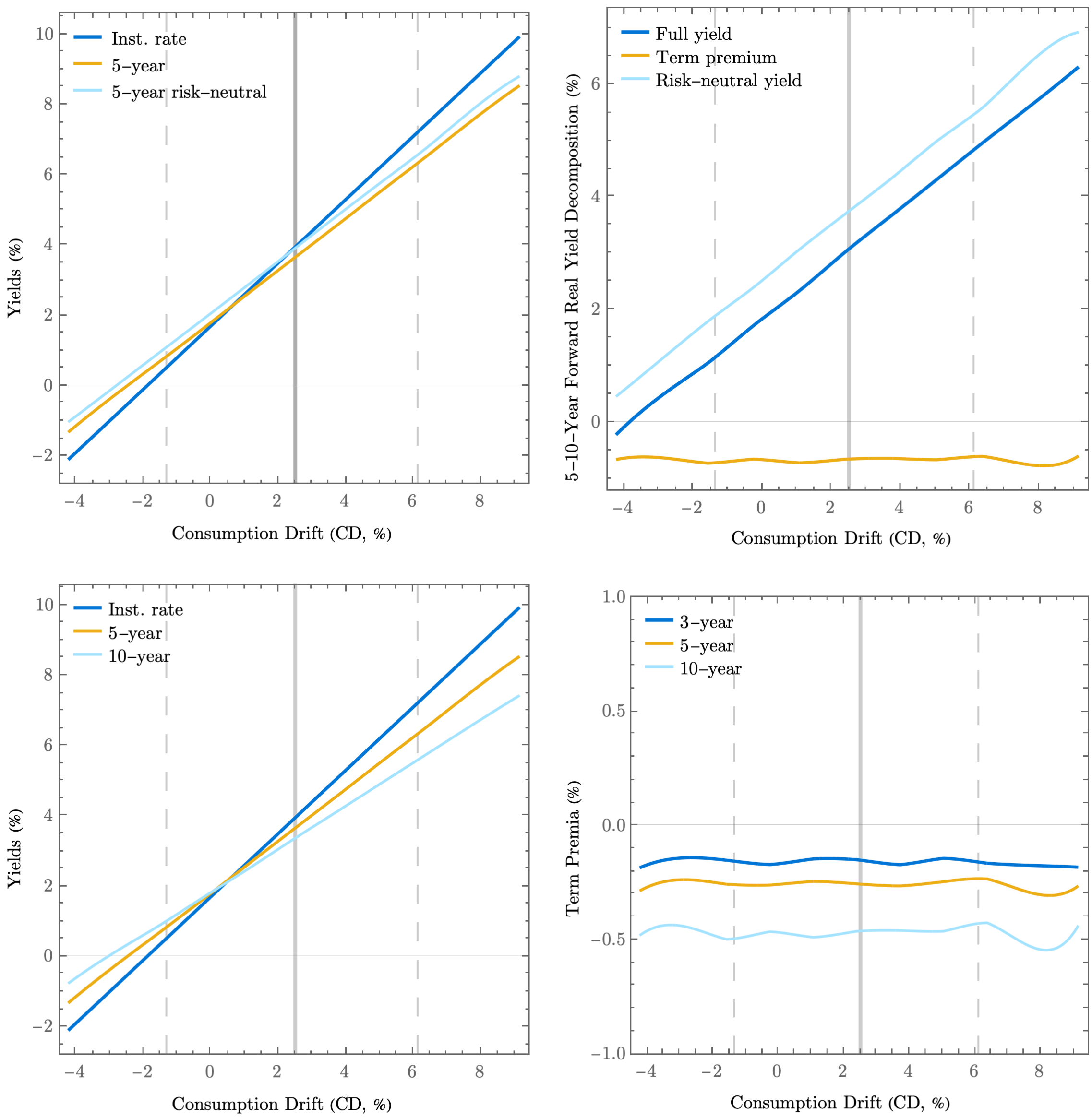


Figure 40: Time-varying CD with RU and high ρ_{cx}.
See Figure 16 for more details about the plots.

(variation overview)

F.28 RU-CD-NCor, $\rho_{cx} < 1$

Term premia increase but they remain negative as in RU term premia are dominated by the term, not including ρ_{cx}.



γ=2, ρ=0.012, μ_{ct}=x_t+0.0252, σ_{ct}=0.02, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=−1, ψ=1.11, approx. order=8

$\mathrm{range}_{\mathrm{low}}(\mathrm{r}_{2,\mathrm{model}})$=0.0484668, $\mathrm{range}_{\mathrm{high}}(\mathrm{r}_{2,\mathrm{model}})$=0.0897518, $\mathrm{range}(\mathrm{r}_{2,\mathrm{data}})$=0.071748.

Figure 41: Time-varying CD with RU and negative ρ_{cx}.
See Figure 16 for more details about the plots.

(variation overview)

F.29 RU-HCD, $\mu_{x0} = 0.05$

Term premia do not change, but yields increase.

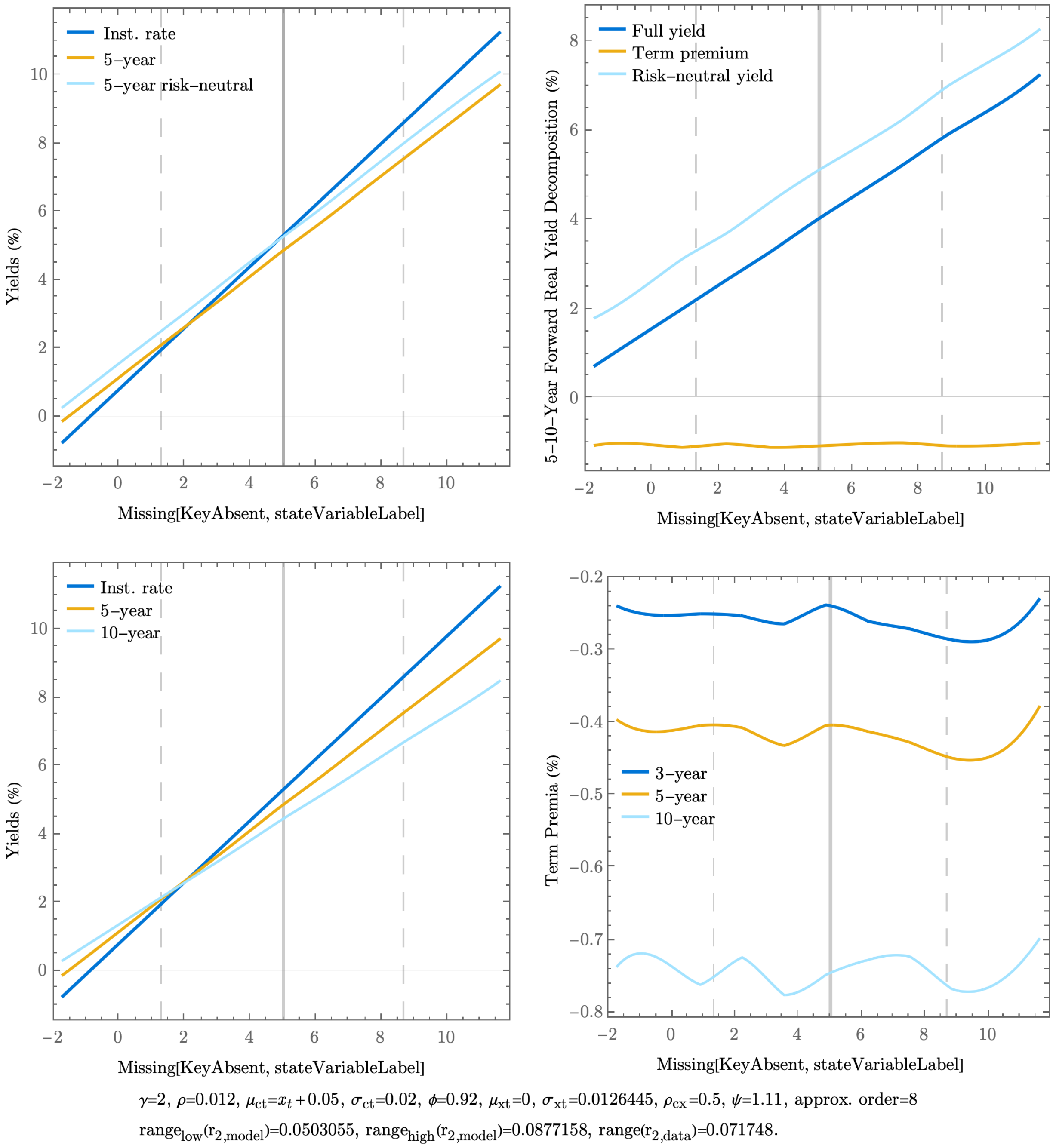
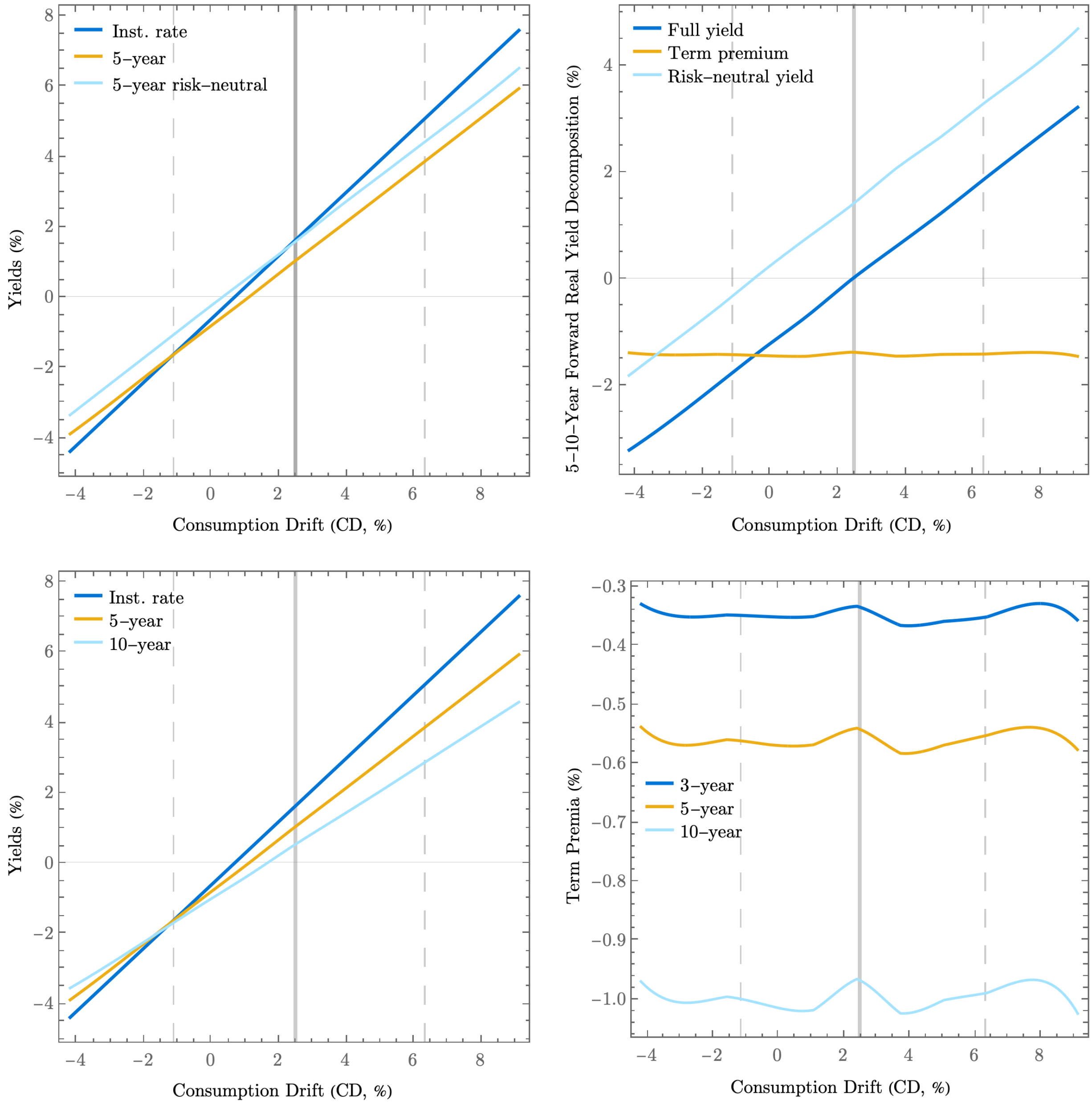


Figure 42: Time-varying and high CD with RU.
See Figure 16 for more details about the plots.

(variation overview)

F.30 RU-CD-HCV, $\sigma_{ct} = 0.08$

Term premia do not change, but yields decrease.



γ=2, ρ=0.012, $\mu_{\mathrm{ct}}=x_t+0.0252$, σ_{ct}=0.08, ϕ=0.92, μ_{xt}=0, σ_{xt}=0.0126445, ρ_{cx}=0.5, ψ=1.11, approx. order=8

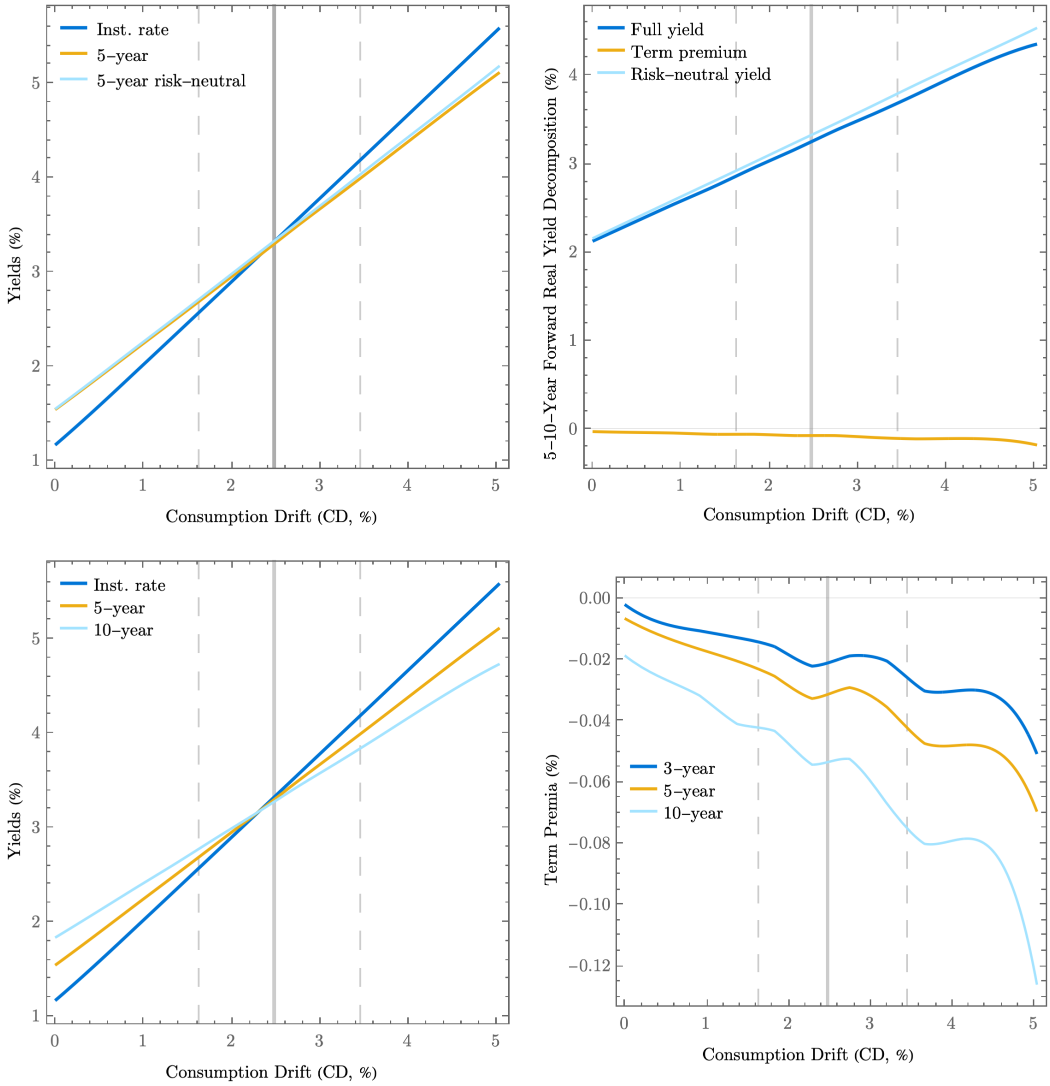
$\mathrm{range}_{\mathrm{low}}(\mathrm{r}_{2,\mathrm{model}})$=0.0492211, $\mathrm{range}_{\mathrm{high}}(\mathrm{r}_{2,\mathrm{model}})$=0.0952063, $\mathrm{range}(\mathrm{r}_{2,\mathrm{data}})$=0.071748.

Figure 43: Time-varying CD with RU and HCV.
See Figure 16 for more details about the plots.

(variation overview)

F.31 RU-CD-Heterosk-PCor

When the state variable is heteroskedastic, term premia become time-varying. Here term premia are quite small, but this could change once a more volatile state variable is introduced. However, term premia are again negative.



γ=2, ρ=0.012, μ_{ct}=0.0252 x_t, σ_{ct}=0.021, ϕ=0.92, μ_{xt}=μ_x, σ_{xt}=0.12 $\sqrt{x_t}$, ρ_{cx}=0.5, ψ=1.11, approx. order=8

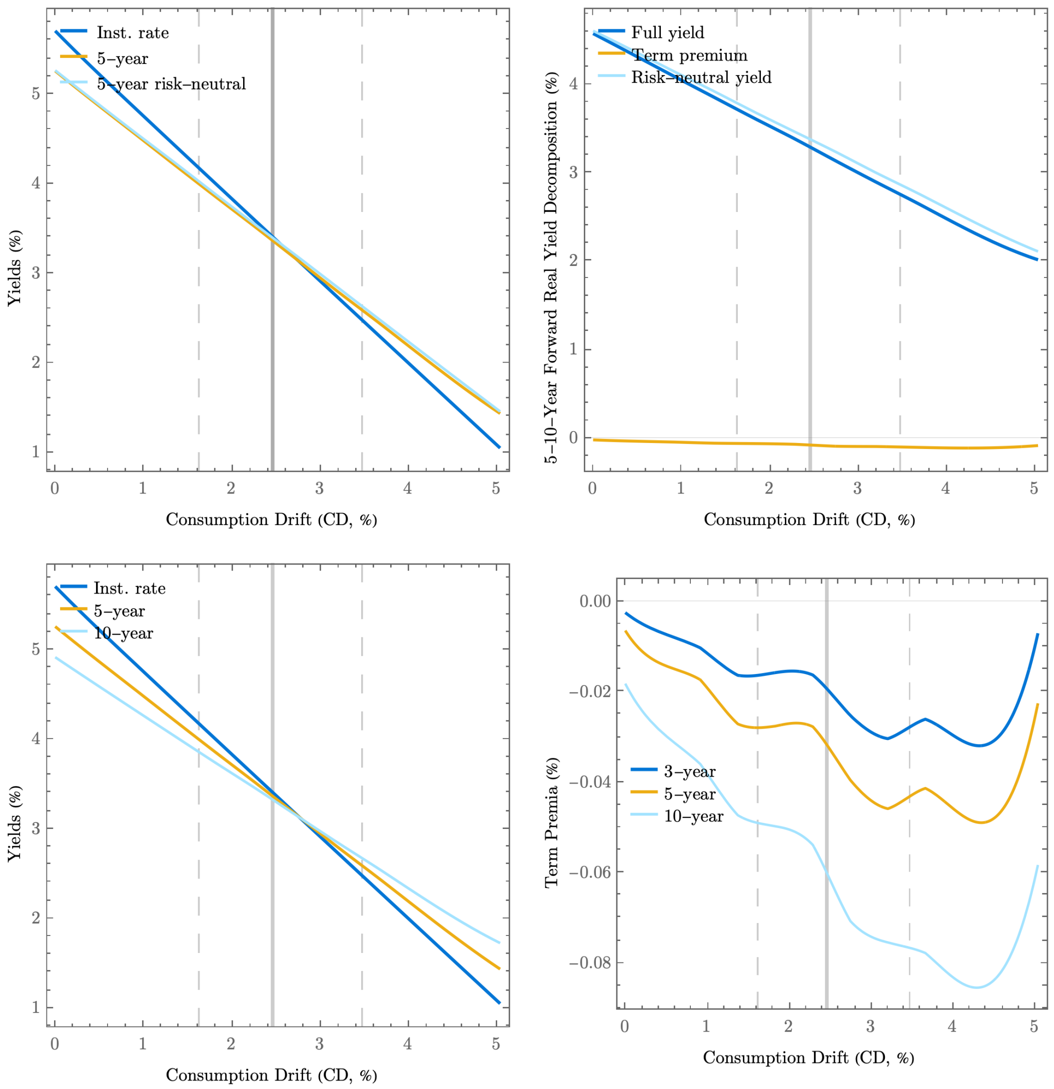
$\mathrm{range}_{\mathrm{low}}(\mathrm{r}_{2,\mathrm{model}})$=0.0115265, $\mathrm{range}_{\mathrm{high}}(\mathrm{r}_{2,\mathrm{model}})$=0.0221465, $\mathrm{range}(\mathrm{r}_{2,\mathrm{data}})$=0.071748.

Figure 44: Time-varying and heteroskedastic CD with RU with positive ρ_{cx}. See Figure 16 for more details about the plots.

(variation overview)

F.32 RU-CD-Heterk-PCor

Despite changing the correlation compared to the previous case term premia are still negative given that the dominant component in function A does not contain ρ_{cx}.



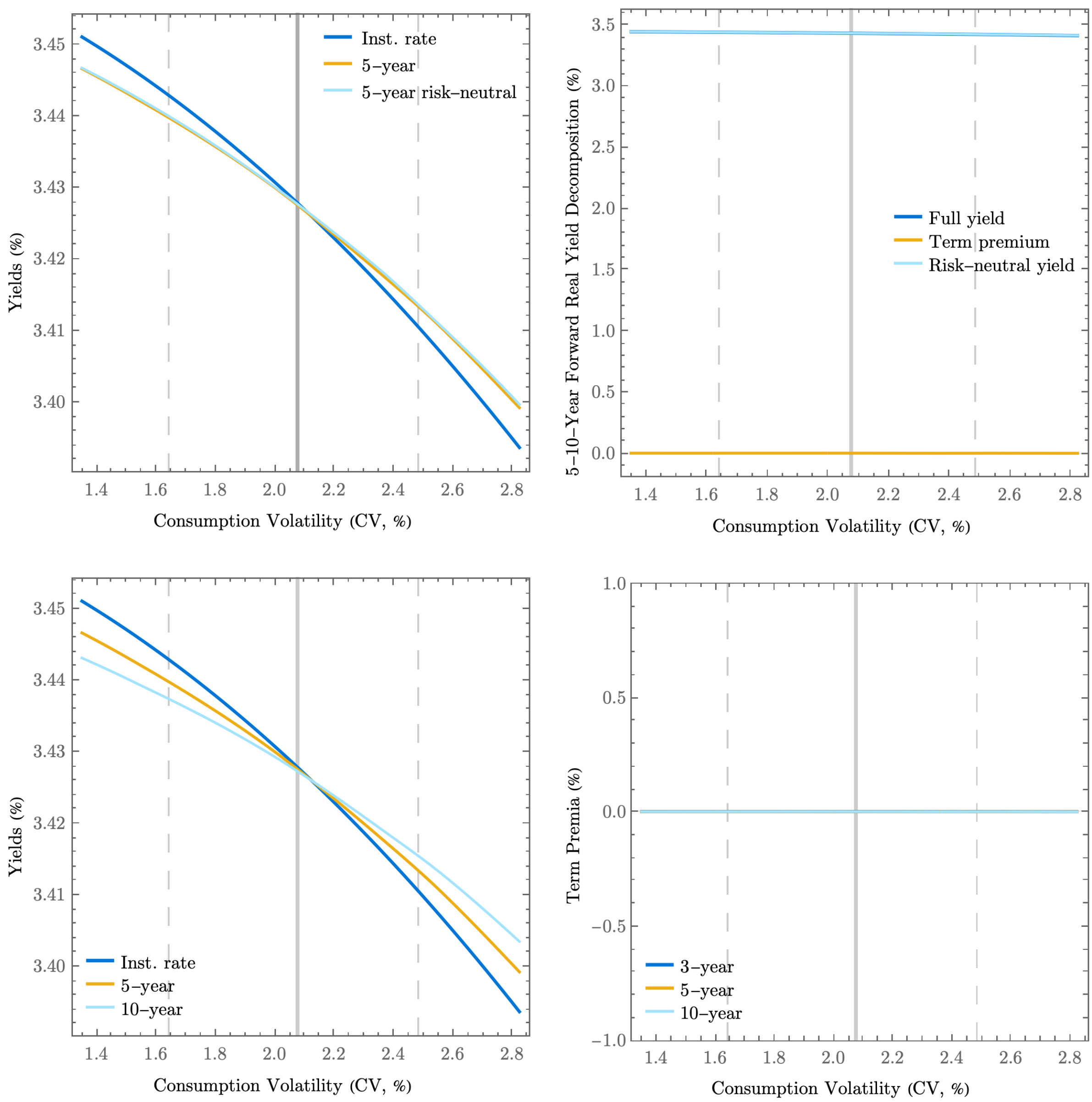
γ=2, ρ=0.012, μ_{ct}=0.0252 $(2-x_t)$, σ_{ct}=0.021, ϕ=0.92, μ_{xt}=μ_x, σ_{xt}=0.12 $\sqrt{x_t}$, ρ_{cx}=−0.5, ψ=1.11, approx. order=8

$\mathrm{range_{low}(r_{2,model})}$=0.0119906, $\mathrm{range_{high}(r_{2,model})}$=0.0230955, $\mathrm{range(r_{2,data})}$=0.071748.

Figure 45: Time-varying and heteroskedastic CD with RU with negative ρ_{cx}. See Figure 16 for more details about the plots.

(variation overview)

F.33 RU-CV, Calibration used in main paper, Figure 5



γ=2, ρ=0.012, μ_{ct}=0.0252, $\sigma_{\text{ct}}=0.021\sqrt{x_t}$, ϕ=0.92, $\mu_{\text{xt}}=\mu_x$, $\sigma_{\text{xt}}=0.13\sqrt{x_t}$, ρ_{cx}=−0.5, ψ=1.11, approx. order=8

$\text{range}_{\text{low}}(\text{r}_{2,\text{model}})$=0.000234865, $\text{range}_{\text{high}}(\text{r}_{2,\text{model}})$=0.000434075, $\text{range}(\text{r}_{2,\text{data}})$=dataRange.

Figure 46: Time-varying CV with RU.
See Figure 16 for more details about the plots.

(variation overview)

F.34 RU-CV-HRA, $\gamma = 6$

The term premia have hardly moved.

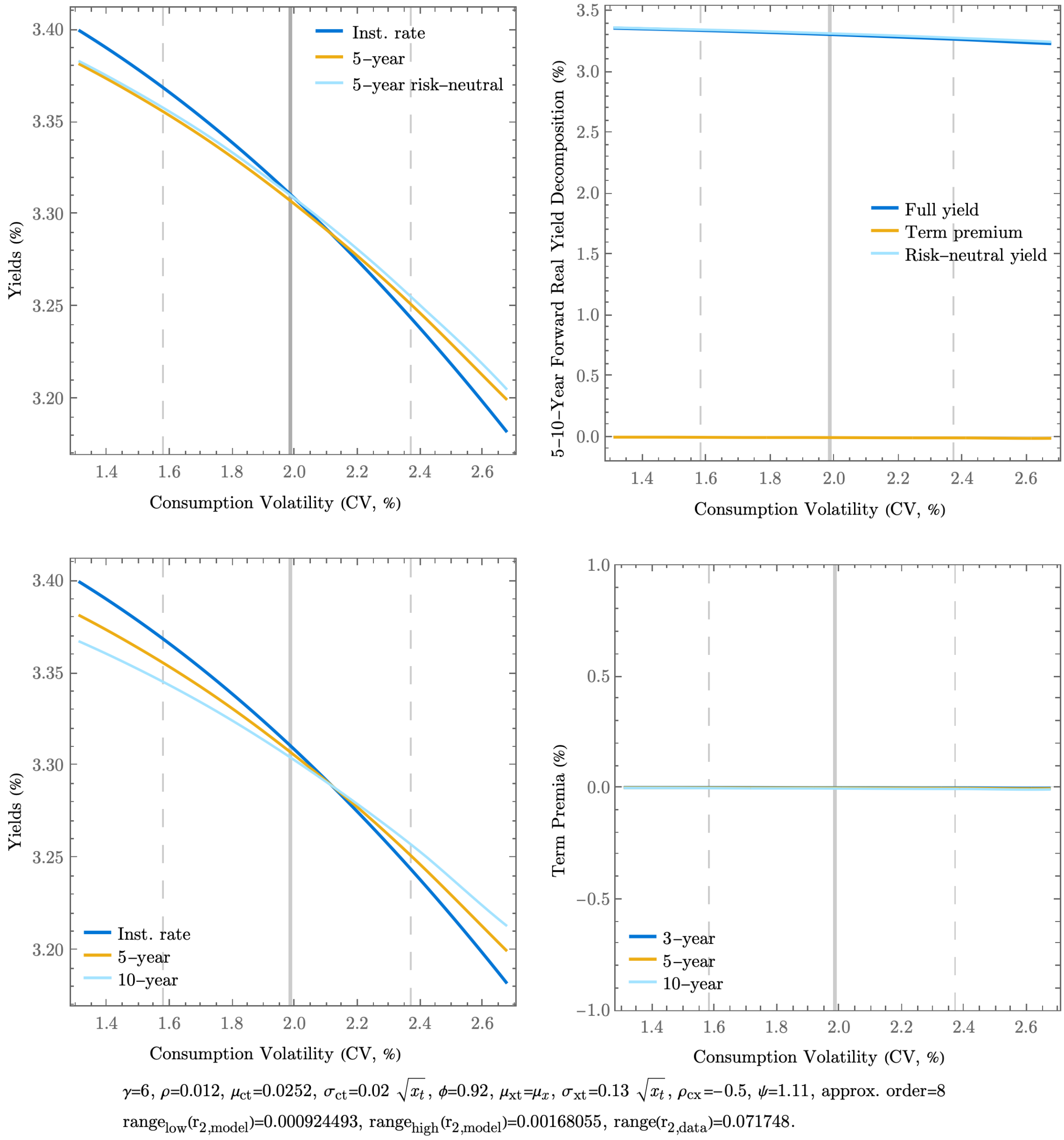
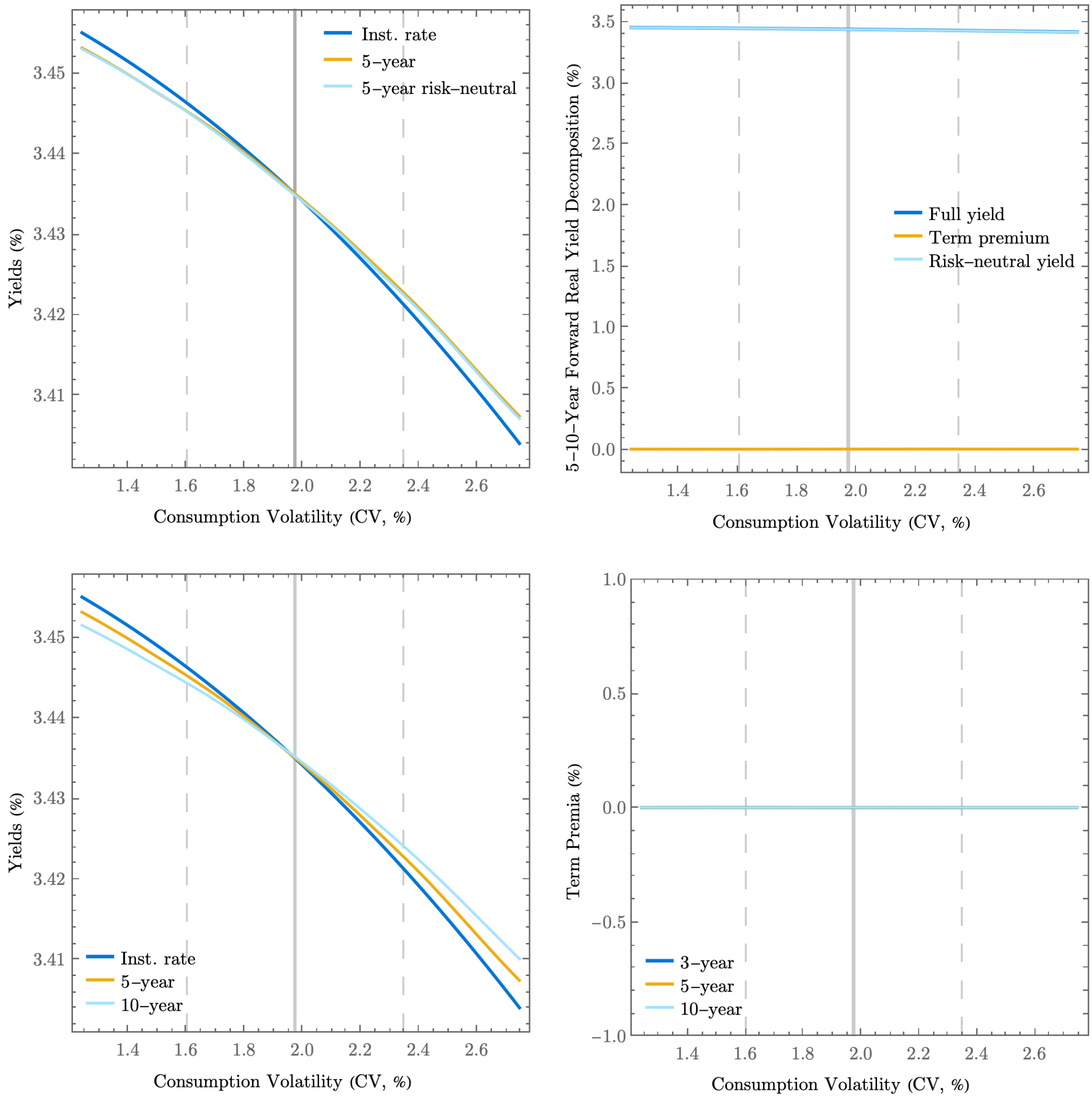


Figure 47: Time-varying CV with RU and high risk aversion.
See Figure 16 for more details about the plots.

(variation overview)

F.35 RU-CV-HP, $\phi = 0.96$

The term premia have hardly moved and curiously the yields have become slightly more variable again.



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02 $\sqrt{x_t}$, ϕ=0.96, μ_{xt}=μ_x, σ_{xt}=0.12 $\sqrt{x_t}$, ρ_{cx}=0.5, ψ=1.11, approx. order=8

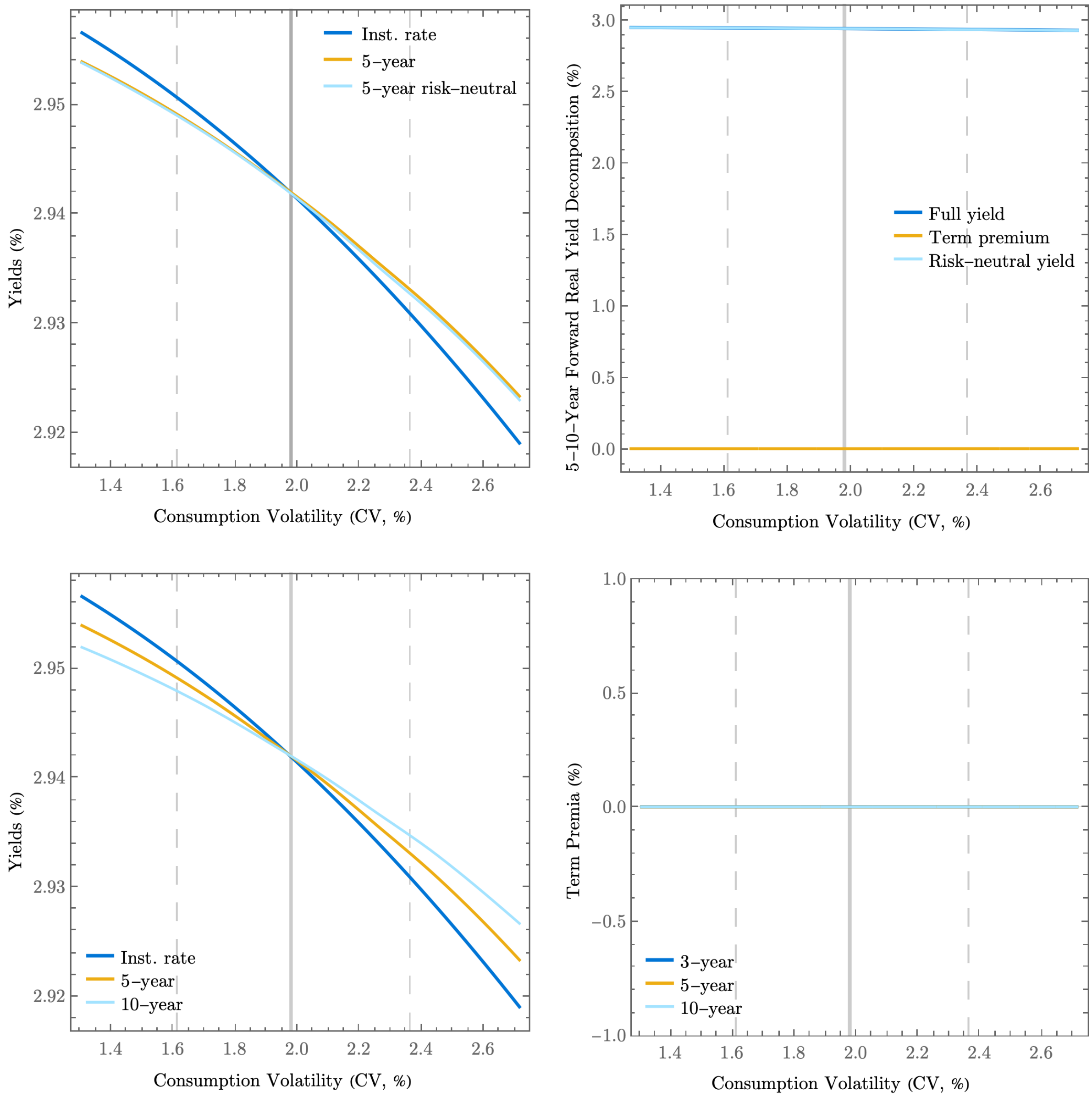
$\mathrm{range}_{\mathrm{low}}(r_{2,\mathrm{model}})$=0.000185272, $\mathrm{range}_{\mathrm{high}}(r_{2,\mathrm{model}})$=0.000369185, $\mathrm{range}(r_{2,\mathrm{data}})$=0.071748.

Figure 48: Time-varying CV with RU and high persistence.
See Figure 16 for more details about the plots.

(variation overview)

F.36 RU-CV-HIES, $\psi = 1.43$

The term premia have hardly moved and the yields have become slightly more variable.



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02 $\sqrt{x_t}$, ϕ=0.92, μ_{xt}=μ_x, σ_{xt}=0.13 $\sqrt{x_t}$, ρ_{cx}=0.5, ψ=1.43, approx. order=8

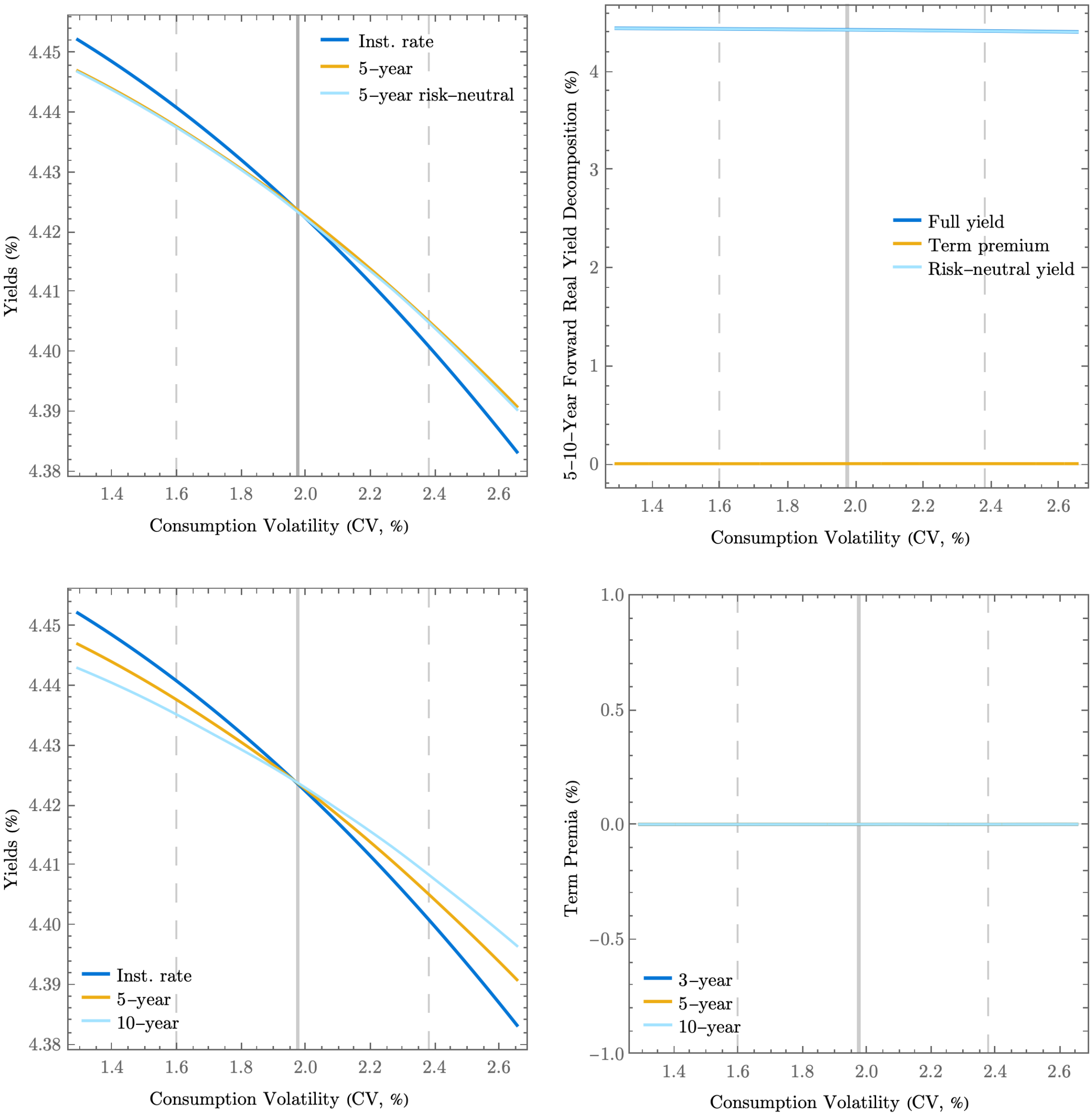
range$_{\mathrm{low}}$(r$_{2,\mathrm{model}}$)=0.000147906, range$_{\mathrm{high}}$(r$_{2,\mathrm{model}}$)=0.000289188, range(r$_{2,\mathrm{data}}$)=0.071748.

Figure 49: Time-varying CV with RU and HIES.
See Figure 16 for more details about the plots.

(variation overview)

F.37 RU-CV-LIES, $\psi = 0.77$

The term premia have hardly moved and curiously the yields have become slightly more variable again.



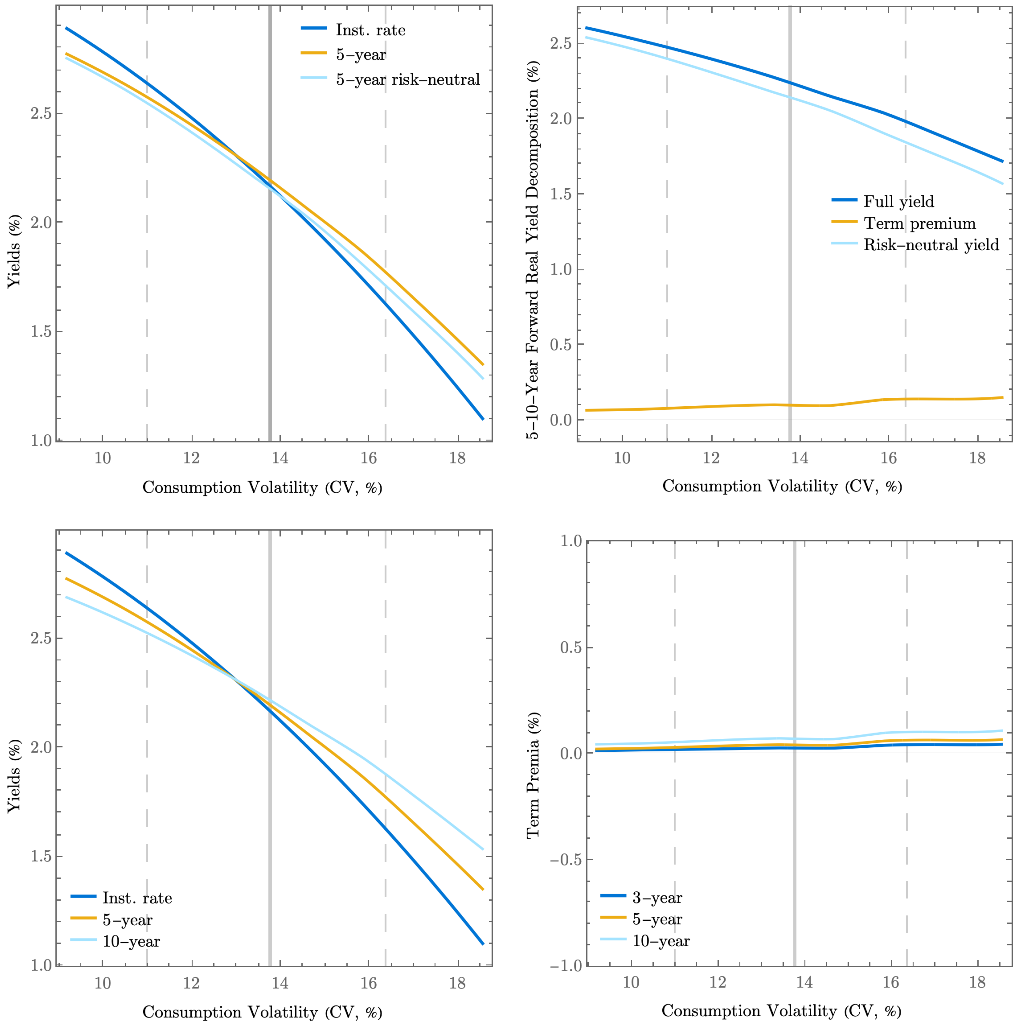
γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.02 $\sqrt{x_t}$, ϕ=0.92, μ_{xt}=μ_x, σ_{xt}=0.13 $\sqrt{x_t}$, ρ_{cx}=0.5, ψ=0.77, approx. order=8
$\text{range}_{\text{low}}(r_{2,\text{model}})$=0.000288099, $\text{range}_{\text{high}}(r_{2,\text{model}})$=0.000549846, $\text{range}(r_{2,\text{data}})$=0.071748.

Figure 50: Time-varying CV with RU and LIES.
See Figure 16 for more details about the plots.

(variation overview)

F.38 RU-HCV-PCor, $\sigma_{c0} = 0.14, \rho_{cx} = 0.5$

Here term premia are positive, which means that the first component of function A that contains ρ_{cx} has become dominant due to the increase in σ_{c1}. Nevertheless, term premia are still smaller than the corresponding term premia in the TSU case, because the second term in function A is still negative.



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.14 $\sqrt{x_t}$, ϕ=0.92, μ_{xt}=μ_x, σ_{xt}=0.13 $\sqrt{x_t}$, ρ_{cx}=0.5, ψ=1.11, approx. order=8

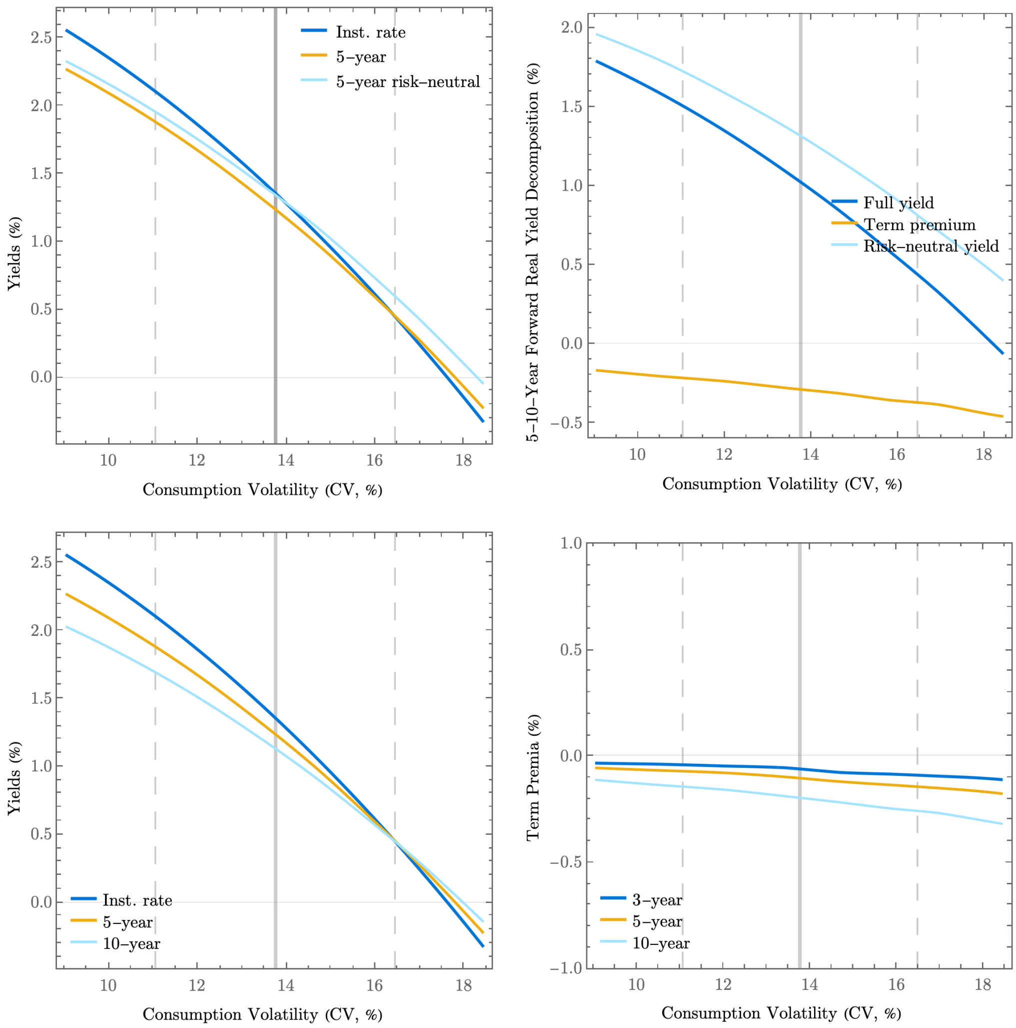
$\mathrm{range_{low}}(\mathrm{r}_{2,\mathrm{model}})$=0.0074128, $\mathrm{range_{high}}(\mathrm{r}_{2,\mathrm{model}})$=0.0146724, $\mathrm{range}(\mathrm{r}_{2,\mathrm{data}})$=0.071748.

Figure 51: Time-varying HCV with RU and PCor.
See Figure 16 for more details about the plots.

(variation overview)

F.39 RU-HCV-NCor, $\sigma_{c0} = 0.14, \rho_{cx} = 0.5$

Here both terms in function A are negative, so term premia are negative. They are also larger in absolute value than the corresponding term premia in RU-CV.



γ=2, ρ=0.012, μ_{ct}=0.0252, σ_{ct}=0.14 $\sqrt{x_t}$, ϕ=0.92, μ_{xt}=μ_x, σ_{xt}=0.13 $\sqrt{x_t}$, ρ_{cx}=−0.5, ψ=1.11, approx. order=8

$\text{range}_{\text{low}}(r_{2,\text{model}})$=0.0120834, $\text{range}_{\text{high}}(r_{2,\text{model}})$=0.0232408, $\text{range}(r_{2,\text{data}})$=0.071748.

Figure 52: Time-varying HCV with RU and NCor.
See Figure 16 for more details about the plots.

(variation overview)

G Deriving the stochastic discount factor

G.1 Derivation of the SDF with TSU

Here the SDE of the SDFis derived, including the case of the habit model. The terms are presented that only apply to the habit model in grey colour. The following is the regular form of the SDF, in which the state variable and log consumption are substituted:

$$\Lambda_t = e^{-\rho t}(e^{c_t}\bar{S}e^{x_t})^{-\gamma} \tag{26}$$

Then, in order to get the SDE form, Ito's Lemma is applied:

$$\begin{aligned}
\mathrm{d}\Lambda_t &= \frac{\partial \Lambda_t}{\partial t}\mathrm{d}t + \frac{\partial \Lambda_t}{\partial c}dc + \underbrace{\frac{\partial \Lambda_t}{\partial x_t}\mathrm{d}x_t}_{\text{habit model}} + \frac{1}{2}\Bigg(\frac{\partial^2 \Lambda_t}{\partial c^2}(\mathrm{d}c_t)^2 + \underbrace{\frac{\partial^2 \Lambda_t}{\partial x^2}(\mathrm{d}x)^2 + \frac{\partial^2 \Lambda_t}{\partial x \partial c_t}\mathrm{d}x_t \mathrm{d}c_t}_{h.m.}\Bigg) \\
&= -\rho\Lambda_t \mathrm{d}t - \gamma\Lambda_t \mathrm{d}c_t - \underbrace{\gamma\Lambda_t \mathrm{d}x_t}_{\text{h.m.}} + \frac{1}{2}\Big(\gamma^2\Lambda_t(\mathrm{d}c_t)^2 + \underbrace{\gamma^2\Lambda_t(\mathrm{d}x_t)^2 + \gamma^2\Lambda_t \mathrm{d}x_t \mathrm{d}c_t}_{h.m.}\Big)
\end{aligned}$$

$$\Rightarrow$$

$$\begin{aligned}
\frac{\mathrm{d}\Lambda_t}{\Lambda_t} &= \Big(-\rho - \gamma\mu_{ct} + \underbrace{\gamma\log(\phi)(\mu_{x0} - x_t)}_{h.m.} + \frac{\gamma^2\sigma_{ct}^2}{2} + \underbrace{\frac{\gamma^2\sigma_{xt}^2}{2} + \gamma^2\rho_{cx}\sigma_{xt}\sigma_{ct}}_{\text{h.m.}}\Big)\mathrm{d}t \\
&\quad - \gamma\sigma_{ct}\mathrm{d}W_{ct} \underbrace{-\gamma\sigma_{xt}\mathrm{d}W_{xt}}_{\text{h.m.}}
\end{aligned} \tag{27}$$

G.2 Derivation of the SDF with RU

As mentioned in the main paper the SDE of the SDF can be derived based on the following expression:

$$\frac{\mathrm{d}\Lambda_t}{\Lambda_t} = F_V(C_t, V_t)\mathrm{d}t + \frac{\mathrm{d}F_C(C_t, V_t)}{F_C(C_t, V_t)} \tag{28}$$

thus, flow utility is a central component of the derivation:

$$F(C_t, V_t) = \frac{\rho}{1 - 1/\psi}\big((1-\gamma)V_t\big)\Bigg(\Big(C_t((1-\gamma)V_t)^{-\frac{1}{1-\gamma}}\Big)^{1-1/\psi} - 1\Bigg) \tag{29}$$

The partial derivative of F with respect to V_t is:

$$F_V(C_t, V_t) = \frac{\rho\left((\gamma-1)\psi + (1-\gamma\psi)\left(C_t(V_t-\gamma V_t)^{\frac{1}{\gamma-1}}\right)^{\frac{\psi-1}{\psi}}\right)}{\psi-1} \tag{30}$$

The partial derivative of F with respect to C_t is:

$$F_C(C_t, V_t) = -\frac{(\gamma-1)\rho V_t\left(C_t(V_t-\gamma V_t)^{\frac{1}{\gamma-1}}\right)^{\frac{\psi-1}{\psi}}}{C_t} \tag{31}$$

As Ito's Lemma is implemented directly using c_t and x_t as independent variables, the following replacements are made in the expressions above:

$$c_t = \log(C_t), \quad V_t = \frac{C_t^{1-\gamma}}{1-\gamma}e^{(1-\gamma)K(x_t)} \Rightarrow K(x_t) = \frac{\log\left(-\frac{C_t^{1-\gamma}}{(\gamma-1)V_t}\right)}{\gamma-1} \tag{32}$$

And after simplification, they become:

$$\begin{aligned} F_V(C_t, V_t) \to G_1(c_t, x_t) =& \frac{\rho(-(1-\gamma\psi)e^{-\frac{(\psi-1)K[x_t]}{\psi}} - \gamma\psi + \psi)}{1-\psi} \\ F_C(C_t, V_t) \to G_2(c_t, x_t) =& \rho e^{(\frac{1}{\psi}-\gamma)K(x_t)-c_t\gamma} \end{aligned} \tag{33}$$

Ito's Lemma is applied on G_2. The partial derivatives are:

$$\begin{aligned} \frac{\partial G_2(c_t,x_t)}{\partial c_t} =& \gamma\rho\left(-e^{\left(\frac{1}{\psi}-\gamma\right)K[x_t]-\gamma c_t}\right) = -\gamma G_2(c_t, x_t) \\ \frac{\partial h(c_t,x_t)}{\partial x_t} =& \rho\left(\frac{1}{\psi}-\gamma\right)K'(x_t)\, e^{\left(\frac{1}{\psi}-\gamma\right)K[x_t]-\gamma c_t} = \left(\frac{1}{\psi}-\gamma\right)K'(x_t)\, G_2(c_t, x_t) \\ \frac{\partial^2 G_2(c_t,x_t)}{\partial c_t^2} =& \gamma^2\rho e^{\left(\frac{1}{\psi}-\gamma\right)K[x_t]-\gamma c_t} = \gamma^2 h(c_t, x_t) \\ \frac{\partial^2 G_2(c_t,x_t)}{\partial x_t^2} =& \frac{\rho(\gamma\psi-1)\left((\gamma\psi-1)K'(x_t)^2 - \psi K''(x_t)\right)e^{\left(\frac{1}{\psi}-\gamma\right)K[x_t]-\gamma c_t}}{\psi^2} \\ =& \frac{(\gamma\psi-1)\left((\gamma\psi-1)K'(x_t)^2 - \psi K''(x_t)\right)}{\psi^2}G_2(c_t, x_t) \\ \frac{\partial G_2(c_t,x_t)}{\partial c_t \partial x_t} =& \frac{\gamma\rho(\gamma\psi-1)K'(x_t)\, e^{\left(\frac{1}{\psi}-\gamma\right)K[x_t]-\gamma c_t}}{\psi} = \frac{\gamma(\gamma\psi-1)K'(x_t)\, G_2(c_t, x_t)}{\psi} \end{aligned} \tag{34}$$

The expressions above should be plugged into the expression:

$$\begin{aligned}\frac{dF_C}{F_C} =&\Bigg(\frac{\partial G_2(c_t,x_t)}{\partial c_t}\mu_{ct} + \frac{\partial G_2(c_t,x_t)}{\partial x_t}\big(-\log(\phi)\big)(\mu_{x0}-x_t)\\ &+\frac{\sigma_{ct}^2}{2}\frac{\partial^2 G_2(c_t,x_t)}{\partial c_t^2} + \frac{\sigma_{xt}^2}{2}\frac{\partial^2 G_2(c_t,x_t)}{\partial x_t^2} + \frac{\rho_{cx}\sigma_{ct}\sigma_{xt}}{2}\frac{\partial^2 G_2(c_t,x_t)}{\partial c_t \partial x_t}\Bigg)\mathrm{d}t\\ &+\frac{\partial G_2(c_t,x_t)}{\partial x_t}\sigma_{xt}\mathrm{d}W_{xt} + \frac{\partial G_2(c_t,x_t)}{\partial c_t}\sigma_{ct}\mathrm{d}W_{ct}\end{aligned} \tag{35}$$

Then everything is plugged into Equation (28) to give the final result:

$$\begin{aligned}\frac{\mathrm{d}\Lambda_t}{\Lambda_t} =&\Bigg(\frac{\gamma(\gamma\psi-1)\rho_{cx}\sigma_{xt}\sigma_{ct}K'(x_t)}{\psi} + \frac{\gamma^2\sigma_{ct}^2}{2} - \gamma\mu_{ct}\\ &+\frac{(\gamma\psi-1)\left(2\psi(\mu_{x0}-x_t)\log(\phi)K'(x_t)+\sigma_{xt}^2\left((\gamma\psi-1)K'(x_t)^2-\psi K''(x_t)\right)\right)}{2\psi^2}\\ &+\frac{\rho\left(-(1-\gamma\psi)e^{-\frac{(\psi-1)K[x_t]}{\psi}}-\gamma\psi+\psi\right)}{1-\psi}\Bigg)\mathrm{d}t - \frac{(\gamma\psi-1)\sigma_{xt}K'(x_t)}{\psi}\mathrm{d}W_{xt} - \gamma\sigma_{ct}\mathrm{d}W_{ct}\end{aligned} \tag{36}$$

References

Abrahams, M., Adrian, T., Crump, R. K., Moench, E. and Yu, R. (2016). Decomposing real and nominal yield curves. *Journal of Monetary Economics*, **84**, 182–200.

Ait-Sahalia, Y., Parker, J. A. and Yogo, M. (2004). Luxury goods and the equity premium. *The Journal of Finance*, **59** (6), 2959–3004.

Angeletos, G.-M., Collard, F. and Dellas, H. (2018). Quantifying confidence. *Econometrica*, **86** (5), 1689–1726.

Backus, D. K., Gregory, A. W. and Zin, S. E. (1989). Risk premiums in the term structure: Evidence from artificial economies. *Journal of Monetary Economics*, **24** (3), 371–399.

Bansal, R. and Shaliastovich, I. (2013). A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets. *The Review of Financial Studies*, **26** (1), 1–33.

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

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, **121** (3), 823–866.

Beechey, M. J. and Wright, J. H. (2009). The high-frequency impact of news on long-term yields and forward rates: Is it real? *Journal of Monetary Economics*, **56** (4), 535–544.

Bekaert, G., Engstrom, E. and Grenadier, S. R. (2010). Stock and bond returns with moody investors. *Journal of Empirical Finance*, **17** (5), 867–894.

Benzoni, L., Collin-Dufresne, P. and Goldstein, R. S. (2011). Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics*, **101** (3), 552–573.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, **107** (2), 205–251.

—, Pflueger, C. and Viceira, L. M. (2020). Macroeconomic drivers of bond and equity risks. *Journal of Political Economy*, **128** (8), 3148–3185.

CARPENTER, S., DEMIRALP, S., IHRIG, J. and KLEE, E. (2015). Analyzing federal reserve asset purchases: From whom does the fed buy? *Journal of Banking & Finance*, **52**, 230–244.

CHEN, Y., COSIMANO, T. F. and HIMONAS, A. A. (2010). Continuous time one-dimensional asset-pricing models with analytic price–dividend functions. *Economic Theory*, **42** (3), 461–503.

—, —, — and KELLY, P. (2009). Asset pricing with long run risk and stochastic differential utility: An analytic approach. *Available at SSRN 1502968*.

COCHRANE, J. H. (2009). *Asset Pricing: Revised edition*. Princeton University Press.

— and PIAZZESI, M. (2005). Bond risk premia. *American Economic Review*, **95** (1), 138–160.

— and — (2009). Decomposing the yield curve. In *AFA 2010 Atlanta Meetings Paper*.

COLLIN-DUFRESNE, P., JOHANNES, M. and LOCHSTOER, L. A. (2016). Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review*, **106** (3), 664–698.

CONSTANTINIDES, G. M. and DUFFIE, D. (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy*, **104** (2), 219–240.

D' AMICO, S., KIM, D. H. and WEI, M. (2018). TIPS from TIPS: The informational content of treasury inflation-protected security prices. *Journal of Financial and Quantitative Analysis*, **53** (1), 395–436.

DEN HAAN, W. J. (1995). The term structure of interest rates in real and monetary economies. *Journal of Economic Dynamics and Control*, **19** (5-7), 909–940.

DONALDSON, J. B., JOHNSEN, T. and MEHRA, R. (1990). On the term structure of interest rates. *Journal of Economic Dynamics and Control*, **14** (3-4), 571–596.

DUFFEE, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, **57** (1), 405–443.

— (2013). Bond pricing and the macroeconomy. *Handbook of the Economics of Finance*, **2**, 907–967.

Duffie, D. and Epstein, L. G. (1992). Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, pp. 353–394.

Ellison, M. and Tischbirek, A. (2021). Beauty contests and the term structure. *Journal of the European Economic Association*, **19** (4), 2234–2282.

Fama, E. F. and Bliss, R. R. (1987). The information in long-maturity forward rates. *The American Economic Review*, pp. 680–692.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics*, **127** (2), 645–700.

Gomez-Cram, R. and Yaron, A. (2021). How important are inflation expectations for the nominal yield curve? *The Review of Financial Studies*, **34** (2), 985–1045.

Gürkaynak, R. S., Sack, B. and Wright, J. H. (2007). The US treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, **54** (8), 2291–2304.

—, — and Wright, J. H. (2010). The TIPS yield curve and inflation compensation. *American Economic Journal: Macroeconomics*, **2** (1), 70–92.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *American Economic Review*, **103** (2), 732–770.

Hsu, A., Li, E. X. and Palomino, F. (2021). Real and nominal equilibrium yield curves. *Management Science*, **67** (2), 1138–1158.

Jappelli, R., Subrahmanyam, M. G. and Pelizzon, L. (2023). Quantitative easing, the repo market, and the term structure of interest rates.

Katagiri, M. (2022). Equilibrium yield curve, the phillips curve, and monetary policy. *Journal of Money, Credit and Banking*, **54** (8), 2235–2272.

Kekre, R., Lenel, M. and Mainardi, F. (2022). Monetary policy, segmentation, and the term structure.

Kliem, M. and Meyer-Gohde, A. (2022). (un) expected monetary policy shocks and term premia. *Journal of Applied Econometrics*, **37** (3), 477–499.

Kraft, H., Seiferling, T. and Seifried, F. T. (2017). Optimal consumption and investment with epstein–zin recursive utility. *Finance and Stochastics*, **21**, 187–226.

Lettau, M. and Wachter, J. A. (2011). The term structures of equity and interest rates. *Journal of Financial Economics*, **101** (1), 90–113.

Malloy, C. J., Moskowitz, T. J. and Vissing-Jørgensen, A. (2009). Long-run stockholder consumption risk and asset returns. *The Journal of Finance*, **64** (6), 2427–2479.

Melissinos, E. (2023). A perturbation solution for the pricing of long-term bonds under recursive utility.

— (2024). Sdfpricing: A julia package for asset pricing based on a stochastic discount factor process.

Pflueger, C. E. and Viceira, L. M. (2016). Return predictability in the treasury market: Real rates, inflation, and liquidity. *Handbook of Fixed-Income Securities*, pp. 191–209.

Piazzesi, M. and Schneider, M. (2006). Equilibrium yield curves. *NBER Macroeconomics Annual*, **21**, 389–442.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics*, **22** (1), 117–131.

Rudebusch, G. D. and Swanson, E. T. (2008). Examining the bond premium puzzle with a DSGE model. *Journal of Monetary Economics*, **55**, S111–S126.

— and — (2012). The bond premium in a DSGE model with long-run real and nominal risks. *American Economic Journal: Macroeconomics*, **4** (1), 105–43.

Savov, A. (2011). Asset pricing with garbage. *The Journal of Finance*, **66** (1), 177–201.

Schneider, A. (2022). Risk-sharing and the term structure of interest rates. *The Journal of Finance*, **77** (4), 2331–2374.

Shiller, R. J. (1979). The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy*, **87** (6), 1190–1219.

Singleton, K. J. (1980). Expectations models of the term structure and implied variance bounds. *Journal of Political Economy*, **88** (6), 1159–1176.

Tsai, J. (2015). Rare disasters and the term structure of interest rates. *Available at SSRN 2669992*.

— and Wachter, J. A. (2018). Pricing long-lived securities in dynamic endowment economies. *Journal of Economic Theory*, **177**, 848–878.

Van Binsbergen, J. H., Fernández-Villaverde, J., Koijen, R. S. and Rubio-Ramírez, J. (2012). The term structure of interest rates in a DSGE model with recursive preferences. *Journal of Monetary Economics*, **59** (7), 634–648.

Vayanos, D. and Vila, J.-L. (2021). A preferred-habitat model of the term structure of interest rates. *Econometrica*, **89** (1), 77–112.

Wachter, J. A. (2006). A consumption-based model of the term structure of interest rates. *Journal of Financial Economics*, **79** (2), 365–399.

— (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, **68** (3), 987–1035.

— and Zhu, Y. (2019). Learning with rare disasters. *Available at SSRN 3407397*.