Ambiguity Aversion and the Term Structure of Interest Rates[*]

Patrick Gagliardini[†]   Paolo Porchia[‡]   Fabio Trojani[§]

JEL CLASSIFICATION: *C*68, *G*12, *G*13.

KEYWORDS: General Equilibrium, Term Structure of Interest Rates, Ambiguity Aversion, Expectations Hypothesis, Campbell-Shiller Regression.

[†]University of Lugano and Swiss Finance Institute, Via Buffi 13, CH-6900 Lugano, Switzerland. E-mail: Patrick.Gagliardini@lu.unisi.ch.

[‡]Swiss Institute for Banking and Finance, University of St. Gallen, Rosenbergstraße 52, CH-9000 St. Gallen, Switzerland. E-mail: Paolo.Porchia@unisg.ch.

[§]University of St. Gallen, University of Lugano and Swiss Finance Institute, Via Buffi 13, CH-6900 Lugano, Switzerland. E-mail: Fabio.Trojani@lu.unisi.ch.

1

Ambiguity Aversion and the Term Structure of Interest Rates

ABSTRACT

This paper studies the term structure implications of a simple structural model in which the representative agent displays ambiguity aversion, modeled by Multiple Priors Recursive Utility. Bond excess returns reflect a premium for ambiguity, which is observationally distinct from the risk premium of affine yield curve models. The ambiguity premium can be large even in the simplest log-utility setting and is also non zero for stochastic factors that have a zero risk premium. A calibrated low-dimensional two-factor model with ambiguity is able to reproduce the deviations from the expectations hypothesis documented in the literature, without modifying in a substantial way the nonlinear mean reversion dynamics of the short interest rate. Moreover, the model does not imply any apparent tradeoff between fitting the first and second moments of the yield curve.

This paper studies the term structure implications of a continuous-time structural model in which the representative agent displays ambiguity (Knightian uncertainty) aversion. First, we characterize the equilibrium in the structural economy with ambiguity aversion to identify the functional form of the equilibrium market prices for risk and ambiguity. Second, we study the theoretical implications for the term structure of interest rates. Third, we calibrate a simple version of the model and study the extent to which ambiguity aversion can reproduce the empirical characteristics of the yield curve. Finally, we investigate how the degree of nonlinearity in the equilibrium short rate dynamics is affected by ambiguity aversion.

We first extend the economy of Cox, Ingersoll, and Ross (1985) by assuming that the representative investor does not know the form of the data-generating process for the underlying production technology. This assumption introduces ambiguity into the economy and is based on a parsimonious one-parameter extension of the standard completely affine Cox, Ingersoll, and Ross (1985) term structure model. We specify investor's aversion to ambiguity by a max-min expected utility optimization problem.[1] Therefore, the excess return of any asset displays a separate premium for ambiguity, which reflects the concern of the representative investor for a misspecification of the technology dynamics. We find that the ambiguity premium can be substantial already for moderate levels of volatility and entails a nonlinear dependence on the underlying state variables.

A rapidly growing literature studies asset prices under ambiguity aversion in dynamic economies. Part of this literature tries to explain the equity premium and interest rate puzzles in a model with a plausible risk aversion parameter. Another part of this literature emphasizes the distinct portfolio behavior of ambiguity-averse investors and the implications for option markets.[2] However, none of these papers studies the relation between ambiguity aversion and the term structure of interest rates. We introduce ambiguity aversion using the tractable setting of Anderson, Hansen, and Sargent (1998, 2003), who measure ambiguity by a maximal constrained discrepancy between a set of likely misspecifications and a fixed reference belief.[3] In this way, we can easily study the impact of ambiguity aversion on the yield curve, in dependence of a single parameter that measures the

3

degree of ambiguity in the economy. The preferences underlying the max-min expected utility problem solved by our representative agent are of the Recursive Multiple Prior Utility type, which implies a 'rectangular' set of relevant likelihoods, in the terminology introduced by Epstein and Schneider (2003), and admits an axiomatic foundation.

The term structure literature has documented several empirical regularities of U.S. Treasury bond yields that are relevant for our work. In a reduced-form term structure setting, Dai and Singleton (2002) and Duffee (2002) emphasize the need for an essentially affine market price of risk specification that relaxes the restrictions of the completely affine models. This specification weakens the link between market price of risk and interest rate volatility for the Gaussian risk factors and improves the empirical predictability patterns for the bond excess returns. Cheridito, Filipovic, and Kimmel (2007) further extend the completely affine specification of the market price of risk also for the volatility factors. In our setting, a non-affine market price of ambiguity relaxes the relation between bond excess returns and volatility, leading to a new class of term structure models. This price of ambiguity is observationally distinct also from the price of risk generated by yield curve models with state-dependent risk aversion and/or money in the utility function, as in Buraschi and Jiltsov (2005, 2007) and Wachter (2006).

Our main findings are as follows. First, in our specification of ambiguity aversion, the market price of ambiguity has a non-affine form. Therefore, the bond excess returns in our model cannot be obtained by any affine yield curve model or another model featuring second-order risk aversion. Moreover, stochastic factors that are not priced in the completely affine models pay an additional premium for ambiguity. Second, a simple two-factor calibrated model with a non-affine market price of ambiguity can reproduce the deviations from the expectations hypothesis documented in the empirical term structure literature. At the same time, it produces a reasonable yields volatility dynamics. Third, the short rate dynamics implied by the model have similar persistence properties as those under the reference belief.[4]

The structure of the paper is as follows. Section 1 presents the opportunity set of our ambiguity-averse representative agent; it defines the set of relevant scenarios in the economy and introduces the max-min expected utility optimization that implies worst case optimal consumption and portfolio policies under ambiguity aversion. Section 2 analyzes equilibrium interest rates, by characterizing the worst case solution in the max-min expected utility optimization, and derives the fundamental differential equation for bond prices. An example of a simplified two-factor model in which the yield curve can be solved in closed form is also derived. Section 3 calibrates a parsimonious two-factor model with ambiguity aversion to bond yield data. It studies the implications of ambiguity aversion for the stylized facts of bond yields and addresses the issue of a potential nonlinearity in the short rate dynamics. Section 4 concludes. All proofs are in the Appendix.

1. Model Setting

The starting point of our analysis is the economy developed by Cox, Ingersoll, and Ross (1985). We introduce ambiguity aversion in this economy to motivate a one-parameter extension of the completely affine market price of risk that can help to explain the well-known yield curve puzzles. On an infinite time horizon, a probability space $(\Omega, \mathcal{F}, P)$ endowed with the filtration $(\mathcal{F}_t)_{t \geq 0}$ supports a $(k+1)-$dimensional standard Brownian motion $Z(t) = [Z_1(t), Z_2(t), \ldots, Z_{k+1}(t)]'$ that generates the uncertainty of the model. We call the probability measure P the 'reference belief.'

1.1 Reference belief

We start deriving the implications of ambiguity aversion for a general form of the dynamics of the technology under the reference belief P. Under the reference belief P, the basic constituents of the opportunity set available to agents are:[5]

1. A locally risk-less bond in zero net supply, with return r.

2. A linear technology Q, producing a physical good that can be either reinvested or consumed. The output rate of the production technology is given by:

$$\frac{dQ}{Q} = \alpha(Y)dt + \sigma(Y)\,dZ. \tag{1}$$

3. k financial assets, in zero net supply, that satisfy the stochastic differential equation:

$$dS = I_S\beta(Y)dt + I_S\,\vartheta(Y)\,dZ, \tag{2}$$

   where $S = [S_1, S_2, \ldots, S_k]'$ is the vector of price processes of these assets and I_S denotes the diagonal matrix $\mathrm{diag}[S_1, S_2, \ldots, S_k]$.[6]

4. k driving state variables $Y = [Y_1, Y_2, \ldots, Y_k]'$ with dynamics:

$$dY = \Lambda(Y)dt + \Xi(Y)\,dZ. \tag{3}$$

The equilibrium in the economy is supported by a single representative agent with a time preference rate $\delta > 0$ and a logarithmic felicity function:[7]

$$U(c,t) = e^{-\delta t}\log(c) \quad ; \quad c > 0 \quad .$$

To focus exclusively on the additional implications of ambiguity aversion for the yield curve, we have chosen a very standard structure of our economy under the reference belief [see Cox, Ingersoll, and Ross (1985)]. In particular, an affine dynamics for dQ and dY implies an affine equilibrium yield curve.

1.2 Model misspecification

The representative agent does not know the data-generating process governing the production technology, the asset returns, and the state variables. She considers a class of probabilistic scenarios, or contaminations, P^h around the reference belief. These contaminations are interpreted as likely specifications for the constituents of the opportunity sets, and are assumed to be absolutely continuous with respect to the reference belief P, as in Anderson, Hansen, and Sargent (1998, 2003). Therefore, contaminations of the reference belief are equivalently described by contaminating drift processes h. Since the scenarios P^h are mutually absolutely continuous, it follows that $Z_h(t) = Z(t) + \int_0^t h(s)ds$ defines a standard Brownian motion process under P^h. Therefore, ambiguity takes the form of a change of drift in the technology dynamics specified with respect to the reference belief.

Aversion to ambiguity arises by assuming that the representative agent is concerned with the worst case scenario in a neighborhood of the reference belief. In order to identify such a neighborhood, we assume that the contaminating drift processes h satisfy the following upper bound:

$$h^{'}h \leq 2\eta, \tag{4}$$

where $\eta \geq 0$ is a fixed parameter. For tractability, we restrict our analysis to the class of Markov Girsanov kernels $h(Y)$ defined by measurable functions $h(\cdot)$. $\mathcal{H}$ denotes the class of admissible Markovian drift contaminations satisfying the bound (4). This family of admissible drift contaminations admits a clear interpretation in terms of a maximal allowed statistical discrepancy between P^h and P [see Anderson, Hansen, and Sargent (1998, 2003)]. Precisely, it implies a maximal bound η on the instantaneous growth rate of the relative entropy between any admissible contaminated model P^h and the reference belief P. It follows that we can interpret η as a measure of the degree of ambiguity in the model: Larger values of η imply a higher ambiguity. Moreover, for $\eta \to 0$, we obtain the standard Cox, Ingersoll, and Ross (1985) setting.

It is convenient to recall that statistically there is little scope for a concern for ambiguity if alternative models are easily detectable. In other words, a reasonable value for the entropy bound η should not imply a too wide discrepancy between the reference model P and alternatives P^h. Anderson, Hansen, and Sargent (2003) propose to assess the adequacy of a given ambiguity bound by considering the probability to erroneously detect an alternative model P^h under the null that model P holds, and vice-versa, using a time-series of finite length. We apply this approach in Section 3 to assess the relevance of the calibrated ambiguity parameter implied by our empirical findings. Finally, note that the bound (4) constrains the instantaneous time-variation of the relative entropy between the two models, not just its global continuation value. Therefore, our specification of ambiguity aversion is based on a rectangular set of priors, using the terminology of Epstein and Schneider (2003), which implies a dynamically consistent preference ordering supported by Multiple Priors Recursive Utility.[8]

1.3 Max-min expected utility

The representative agent finances her consumption process $c(t)$ by trading continuously in the financial assets at the equilibrium prices. If we denote by Σ the $(k+1) \times (k+1)$ diffusion matrix of the available opportunity set,

$$\Sigma(Y) = \begin{bmatrix} \sigma(Y) \\ \vartheta(Y) \end{bmatrix}, \tag{5}$$

then the usual dynamic budget constraint, coupled with the appropriate integrability conditions,[9] implies feasibility of consumption plans:

$$\frac{dW(t)}{W(t)} = \left[\omega(t)\left(\alpha(t) - r(t)\right) + v(t)\left(\beta(t) - r(t)\overline{1}_k\right) + \left(r(t) - \frac{c(t)}{W(t)}\right)\right] dt$$

$$+ \pi(t)'\Sigma(t)\left[dZ(t) + h(t)\,dt\right], \tag{6}$$

where $\pi = [\omega \quad v]' \in \mathbb{R} \times \mathbb{R}^k$ contains the portfolio weights ω and v' invested in the technology and the financial assets, respectively, $\overline{1}_k$ is a $k-$dimensional vector of ones, and $W(t)$ is the financial wealth of the representative agent at time t. In order to prevent arbitrage opportunities, we follow Dybvig and Huang (1988) and assume that $W(t) \geq 0$ for every $t \geq 0$. The ambiguity-averse representative investor solves the max-min expected utility program:

$$J(x,y) = \sup_{c,\pi} \quad \inf_{h \in \mathcal{H}} \mathbb{E}^h \left[\int_0^\infty e^{-\delta s} \log(c(s)) ds \right],$$ (7)

$$\text{s.t.} \quad (6),$$

where $W(0) = x$, $Y(0) = y$ and $\mathbb{E}^h[\cdot]$ denotes expectations under measure P^h. In a Cox, Ingersoll, and Ross (1985) economy, financial securities are in zero net supply. Therefore, their expected returns are shadow prices for the constraint to hold a null portfolio weight on these securities. We have the following definition of equilibrium.

Definition 1. *An equilibrium is a vector (c^*, h^*, r^*, β^*) of a consumption policy, a model contamination, interest rate, and financial assets return processes, such that:*

1) *The equilibrium consumption policy c^* and the drift contamination h^* are optimal according to the following preference ordering representation:*

$$\inf_{h \in \mathcal{H}} \mathbb{E}^h \left[\int_0^\infty e^{-\delta s} \log(c(s)) ds \right].$$ (8)

2) *Optimal consumption is financed by a trading strategy according to which wealth is totally invested in the technology:*

$$\pi = [\omega \quad v]' \equiv [1 \quad \mathbf{0}]',$$ (9)

where $\mathbf{0}$ is a k-dimensional row vector of zeros.

9

2. Characterization of Equilibrium and Pricing

This section studies the equilibrium in our model with ambiguity aversion, the functional form of the resulting market price of ambiguity, and its implications for the yield curve.

2.1 Equilibrium value function

In equilibrium, the value function (7) is evaluated at the market-clearing values of the short interest rate and the returns on financial assets. This value function depends on both the model selection in the corresponding max–min expected utility problem and the optimality conditions evaluated at equilibrium prices. We can postpone the selection of the optimal contaminating drift h^* to the determination of the dependence of the interest rate and the excess returns of financial assets on any admissible contaminating drift $h \in \mathcal{H}$. Therefore, we interchange the order of the maximization and the minimization in problem (7) and observe that the resulting innermost program is a standard optimization, in which the equilibrium conditions are easily handled by standard methods.[10] Proposition 1 exploits this fact to characterize the optimal contaminating drift h^* in our economy.

Proposition 1. *In equilibrium, the value function of the ambiguity averse representative investor is given by:*

$$J(x,y) = -\frac{1}{\delta} + \frac{\log(\delta x)}{\delta} + \frac{1}{\delta}V(y), \tag{10}$$

where:

$$V(y) = \inf_{h \in \mathcal{H}} \mathbb{E}^h \left[\int_0^\infty e^{-\delta t} \left(\alpha(t) - \frac{1}{2}\sigma(t)\sigma(t)' + \sigma(t)h(t) \right) dt \right], \tag{11}$$

subject to:

$$dY = \Lambda(Y)\,dt + \Xi(Y)\,(dZ + h\,dt).$$

The equilibrium contaminating drift that solves the model selection problem in the max-min optimization (7) is given by:

$$h^* = -\sqrt{2\eta}\, \frac{\Xi' V_Y + \sigma'}{\sqrt{(\Xi' V_Y + \sigma')'(\Xi' V_Y + \sigma')}},$$ (12)

where V_Y denotes the gradient of V with respect to the state variables Y. The value function V solves the following Hamilton-Jacobi-Bellman (HJB) equation:

$$V_Y' \Lambda + \frac{1}{2}\text{trace}\left[\Xi' V_{YY} \Xi\right] - \sqrt{2\eta}\sqrt{(\Xi' V_Y + \sigma')'(\Xi' V_Y + \sigma')} + \alpha - \frac{1}{2}\sigma\sigma' - \delta V = 0,$$ (13)

where V_{YY} is the Hessian matrix of V with respect to the state variable Y.

The value function J of the ambiguity-averse representative investor depends on value function V, which solves the HJB equation (13). In the definition of V, the contaminating drift h affects the state variables dynamics by shifting probabilities over the space of sample paths in an absolutely continuous fashion. The expectation inside Equation (11) is defined over a discounted "running cost" $e^{-\delta t} l(t)$, say, where:

$$l(t) = \alpha(t) - \frac{1}{2}\sigma(t)\sigma(t)' + \sigma(t)h(t).$$ (14)

The difference of the first two terms in the right-hand side reflects the impact of the standard risk-return tradeoff on the indirect utility of a representative agent with logarithmic utility function in a Cox, Ingersoll, and Ross (1985) production economy. This difference depends on the variance $\sigma(t)\sigma(t)'$ of the production technology. Therefore, it affects the indirect utility of the representative agent by means of the well-known second-order risk-aversion effect. The third term in the right-hand side of (14) is proportional to the volatility of the production technology, weighted by the contaminating drift h. If the optimal control h^* is constant, the contribution of this term is proportional to $\sigma(t)$ and impacts on the indirect utility of the representative agent in a way that

is observationally equivalent to a first-order risk-aversion effect. It is easy to show that this feature arises in our economy when the admissible drift-contaminations $h \in \mathcal{H}$ are one-dimensional. For a general multi-factor technology with stochastic volatility, the optimal worst case drift contamination is state-dependent. In this case, ambiguity aversion impacts on the indirect utility of the representative agent in a way that is different from the effect of both first- and second-order risk-aversion.[11]

2.2 Equilibrium interest rate and market price of ambiguity

The equilibrium interest rate and the market prices of risk and ambiguity follow as a direct Corollary to Proposition 1.

Corollary 1. *The market price of risk and ambiguity in the economy with ambiguity aversion is given by:*

$$\lambda = \sigma' + \sqrt{2\eta} \frac{\Xi' V_Y + \sigma'}{\sqrt{(\Xi' V_Y + \sigma')'(\Xi' V_Y + \sigma')}} =: \lambda_R + \lambda_A, \tag{15}$$

and the equilibrium interest rate is:

$$r = \alpha - \sigma\lambda. \tag{16}$$

The market price of risk and ambiguity is the sum of the market price of risk $\lambda_R = \sigma'$ and the market price of ambiguity:

$$\lambda_A = \sqrt{2\eta} \frac{\Xi' V_Y + \sigma'}{\sqrt{(\Xi' V_Y + \sigma')'(\Xi' V_Y + \sigma')}}. \tag{17}$$

λ_R takes the standard completely affine form, in which risk compensation per unit of risk is identical to the volatility of the production technology. Without ambiguity ($\eta = 0$), this feature implies a strong relation between excess returns and the volatility of financial assets in the economy, which

is at odds with the stylized facts of the U.S. Treasury yields. Dai and Singleton (2002) and Duffee (2002) emphasize the need of a more flexible market price of risk specification, linked to factors that might affect the slope of the yield curve, in addition to the volatility. To this end, Duffee (2002) proposes the essentially affine specification of the market price of risk, which breaks the link between the market price of risk and the volatility of Gaussian factors. With the same purpose, Cheridito, Filipovic, and Kimmel (2007) have further extended the market price of risk specification for the stochastic volatility factors in reduced–form models. In our model, the market price of ambiguity λ_A breaks the connection between excess returns and volatility, but its nonlinear form implies a yield curve that is not affine. λ_A depends on Ξ, the volatility matrix of all state variables, and the marginal utility V_Y of the representative investor with respect to the state variables. This is the reason why bonds excess returns can be associated with other state variables than volatility. Moreover, the components of the product $\Xi' V_Y$ can change sign over time. Therefore, bond excess returns can take both positive and negative values, consistent with the empirical evidence. Finally, since the component of bond excess returns due to ambiguity is proportional to volatility, while the component due to risk is proportional to variance, the former dominates the latter when the aggregate risk in the economy is low. [12]

The dependence of the market price of ambiguity on the term $\Xi' V_Y$ is due to the non-myopic portfolio behavior of the ambiguity-averse representative agent, even with logarithmic utility. Intuitively, the representative investor hedges her portfolio against future changes in the worst case opportunity set, which are in a one-to-one relation to the realizations of the worst case drift distortion h^*. This hedging motive is different from the standard hedging motive against the risk of a change in the opportunity set, as perceived under the reference belief, which is known to vanish for log-utility investors. It follows that state variables uncorrelated with the technology process can pay a premium for ambiguity, even if their risk would not be priced in a standard Cox, Ingersoll, and Ross (1985) economy.

2.3 Term structure of interest rates

Given the expression for the market price of risk and ambiguity in Corollary 1, we can now price any interest rate derivative by standard arbitrage arguments. The change of drift ϕ_Y from the reference belief P to the risk-neutral probability measure $\mathbb{Q}$ is given by:

$$\phi_Y = \Xi\lambda = \Xi\left[\sigma' + \sqrt{2\eta}\left(\frac{\Xi'V_Y + \sigma'}{\sqrt{(\Xi'V_Y + \sigma')'(\Xi'V_Y + \sigma')}}\right)\right], \tag{18}$$

where the value function V solves the HBJ equation (13). Given the drift change ϕ_Y, the price of a contingent claim with maturity T and paying off at a rate $\Psi(Y,t)$, $t \leq T$, is characterized by the following fundamental partial differential equation.

Proposition 2. *The price at time t, $F(Y,t)$, of a contingent claim with instantaneous pay-off $\Psi(Y,t)$, $t \leq T$, satisfies the partial differential equation:*

$$\frac{1}{2}\text{trace}\left(\Xi\Xi'\frac{\partial^2 F}{\partial Y \partial Y'}\right) + (\Lambda - \phi_Y)'\frac{\partial F}{\partial Y} - rF + \frac{\partial F}{\partial t} = -\Psi, \tag{19}$$

with boundary condition $F(Y,T) = \Psi(Y,T)$, where r is the equilibrium short rate given in Corollary 1 and ϕ_Y is the risk-neutral drift change defined in Equation (18).

Compared to the completely affine setting, ambiguity aversion alters the fundamental pricing equation through the different equilibrium interest rate r and the corresponding change of drift ϕ_Y. Therefore, the Feynman-Kac theorem gives the usual probabilistic representation of the derivative price:

$$F(Y,t) = \mathbb{E}_{\mathbb{Q}}\left[\int_t^T e^{-\int_t^s r(u)du}\Psi(Y(s),s)ds + e^{-\int_t^T r(s)ds}\Psi(Y(T),T)\,\middle|\,\mathcal{F}_t\right], \tag{20}$$

where $\mathbb{E}_{\mathbb{Q}}[\cdot]$ denotes the expectation with respect to the risk-neutral probability measure $\mathbb{Q}$ under ambiguity aversion. A major difference with respect to the pricing problem of a completely affine

14

Cox, Ingersoll, and Ross (1985) economy emerges, because the equilibrium interest rate r and the drift change ϕ_Y are preference-dependent parameters when ambiguity aversion is present. They are determined by the value function gradient V_Y appearing in the formulas of Corollary 1. Therefore, the pricing problem cannot be separated from the equilibrium computation of the market price of ambiguity λ_A in Equation (17). As a consequence, in the economy with ambiguity aversion, we have to solve the system of partial differential Equations (13) and (19) in order to compute the equilibrium yield curve.[13]

2.4 Closed-form solutions in a simple model setting

We first discuss a simplified two-factor model in which the representative investor exhibits ambiguity only over one of the state variables driving the reference belief. This assumption implies a more restrictive form of the ambiguity premium, but yields closed-form expressions for the term structure of interest rates. Therefore, this model provides a basic intuition for the role of ambiguity aversion in the context of the yield curve. In addition, it clarifies the limitations of an economy in which the structure of ambiguity is independent across the relevant state variables. A more general economy in which ambiguity aversion affects the subjective dynamics of all states variables is analyzed in Section 3, where we calibrate the model to U.S. Treasury yields data.

2.4.1 State variables process. The reference belief dynamics satisfies a two-factor Longstaff and Schwartz (1992) model. Therefore, under any model contamination P^h, the production technology dynamics can be written as:

$$\frac{dQ}{Q} = (g_1 Y_1 + g_2 Y_2)\,dt + l\sqrt{Y_2}\left(\sqrt{1-\rho^2}\,(dZ_1 + h_1 dt) + \rho\,(dZ_3 + h_3 dt)\right), \tag{21}$$

$$dY_1 = (a_1 + m_1 Y_1)\,dt + n_1\sqrt{Y_1}\,(dZ_2 + h_2 dt), \tag{22}$$

$$dY_2 = (a_2 + m_2 Y_2)\,dt + n_2\sqrt{Y_2}\,(dZ_3 + h_3 dt), \tag{23}$$

where $Z = [Z_1, Z_2, Z_3]'$ is a three-dimensional standard Brownian motion, and g_i, a_i, m_i, n_i, $i = 1, 2$, and l are scalar parameters. In this model, the instantaneous correlation between dQ/Q and dY_1 is zero, but the correlation between dQ/Q and dY_2 is allowed to be nonzero, and is given by the parameter ρ. It follows that the variance of the technology return is the only state variable, in addition to the technology itself, which pays a non-zero risk premium in this economy. For illustration purposes and for analytical tractability, we make the following assumption, which implies that the representative investor perceives ambiguity only about the shocks in the dynamics of the conditional variance of the technology.

Assumption 1. *The admissible contaminating drift process $h = (h_1, h_2, h_3)'$ is restricted to be of the form $h = (0, 0, h_3)'$. Moreover, the parameter constraint $a_2 = n_2^2/4$ holds.*

The parameter constraint on a_2 in Assumption 1 allows us to obtain closed-form solutions for the term structure of interest rates.

2.4.1 Ambiguity premium. Under Assumption 1, it is easy to show from Proposition 1 that the optimal contaminating drift is $h^* = (0, 0, -\sqrt{2\eta})'$. The market price of risk and ambiguity vector is then immediately inferred from Corollary 1:

$$\lambda = l\sqrt{Y_2} \underbrace{\begin{pmatrix} \sqrt{1-\rho^2} \\ 0 \\ \rho \end{pmatrix}}_{\lambda_R} + \sqrt{2\eta} \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\lambda_A}. \tag{24}$$

The market price of risk λ_R is completely affine, since it is proportional to the volatility of the production technology. The market price of ambiguity λ_A is constant and loads only on the Brownian shock dZ_3 in the volatility of the production technology. This is intuitive since, by Assumption 1, ambiguity aversion does not impact the other shocks in the economy. It follows that even when the market price of risk for the shock dZ_3 in the technology volatility is zero ($\rho = 0$), the market price of ambiguity for these shocks is positive. Moreover, the market price of ambiguity does not

vanish when the production technology risk vanishes ($Y_2 \downarrow 0$). Therefore, as discussed in Subsection 2.2 for the general setting, ambiguity premia tend to dominate risk premia precisely when the aggregate risk in the economy is low. The observationally different form of the market price of ambiguity λ_A follows from our specification of ambiguity aversion in the max-min problem (7), which gives rise to an inherently different behavior from the one generated by the standard risk aversion. Standard risk aversion behavior in the Cox, Ingersoll, and Ross (1985) model implies a concern for variance-covariance risk, i.e., second-order risk aversion, which generates equilibrium expected excess returns that are proportional to variances and covariances of the risk factors. In contrast, in presence of factors with stochastic volatility, our model of ambiguity aversion generates a concern for small risks that does not fit neither into this framework, nor in other essentially affine or extended affine model settings.[14]

2.4.3 Term structure and bond excess returns.

Under Assumption 1, a remarkable feature of the model with ambiguity is that the differential Equations (13) and (19) needed to compute the yield curve are both solvable in closed-form. The price $P(t,T)$ of a zero bond of maturity T is given by:

$$P(t,T) = \exp\left(A(t,T,\eta) + B(t,T)Y_2 + \sqrt{2\eta}C(t,T)\sqrt{Y_2} + D(t,T)Y_1\right),\qquad(25)$$

with functions $B(t,T)$, $C(t,T)$, and $D(t,T)$ admitting closed-form expressions.[15] The nonlinear dependence of $\log P(t,T)$ on the state variable Y_2 emphasizes the non-affine structure of the model. Using the explicit expression for the change of drift ϕ_Y in Equation (18),

$$\phi_Y = \Xi\lambda = \begin{pmatrix} 0 & \sqrt{Y_1} & 0 \\ 0 & 0 & \sqrt{Y_2} \end{pmatrix} \begin{pmatrix} l\sqrt{1-\rho^2}\sqrt{Y_2} \\ 0 \\ l\rho\sqrt{Y_2} + \sqrt{2\eta} \end{pmatrix} = \begin{pmatrix} 0 \\ l\rho Y_2 + \sqrt{2\eta Y_2} \end{pmatrix},\qquad(26)$$

17

we can also easily compute the instantaneous expected excess return on the zero bond:

$$\frac{1}{dt}\mathbb{E}_t\left[\frac{dP(t,T)}{P(t,T)}\right] - r = \frac{P_Y(t,T)}{P(t,T)}\phi_Y$$

$$= \left(B(t,T) + \frac{C(t,T)}{2}\sqrt{\frac{2\eta}{Y_2}}\right)\left(l\rho Y_2 + \sqrt{2\eta Y_2}\right)$$

$$= l\rho B(t,T)Y_2 + \sqrt{2\eta}\left(B(t,T) + l\rho\frac{C(t,T)}{2}\right)\sqrt{Y_2} + \eta C(t,T), \quad (27)$$

where $P_Y(t,T)$ denotes the gradient of the zero coupon price with respect to the state variables. It follows that when $\eta > 0$, the excess returns on bonds are non-affine functions of the state variable Y_2. This simple extension of the affine economy implies a richer behavior of the term structure of bond excess returns, especially for states of moderate volatility risk. More generally, when Assumption 1 is not satisfied, bond excess returns are non-affine functions of all state variables in the model. This additional flexibility of the model is potentially useful for explaining the stylized facts of bond returns. Empirical research has accumulated substantial evidence that bond excess returns are predictable by term structure variables, such as the term structure slope, the spot-forward spread, or a linear combination of forward rates [Fama and Bliss (1987); Campbell and Shiller (1991); and Cochrane and Piazzesi (2005)]. In the next section, we investigate if ambiguity aversion can help explain this empirical evidence within a simple low-dimensional two-factor model.

3. Model Calibration and Empirical Analysis

In this section, we investigate empirically the ability of the non-affine market price of ambiguity in our model to explain some well-known yield curve puzzles. To this end, we calibrate a two-factor model with ambiguity aversion to U.S. Treasury yields data and investigate the empirical yield curve implications of this setting. Using low-dimensional two-factor dynamics to explain the stylized facts of the yield curve is a difficult task. Dai and Singleton (2003), for example, show that three-factor essentially affine models have difficulties in matching at the same time the first

and the second moments of yields. In particular, a completely affine model with three-factor CIR dynamics fails completely in fitting the predictability of bond returns. Therefore, we show that our one-parameter extension of the completely affine model class can help in reproducing some of these stylized facts.

The relevant dynamics for the state variables under any admissible drift contamination $h \in \mathcal{H}$, defined by the bound (4), is given in Equations (21)-(23) without imposing Assumption 1. In contrast to the model presented in Subsection 2.4, this specification does not admit closed-form value function and yield curve solutions for the differential equations (13) and (19).

3.1 Calibration of the model

We calibrate our model using interest rate data for the sample period between January 1960 and December 2000. The interest rate data from January 1960 to February 1991 are obtained from the McCulloch (1990) and Kwon (1992) data-set, which was extended with the methodology described in Buraschi and Jiltsov (2007) until December 2000.[16] The fact that both solutions of the differential equations (13) and (19) are not given in closed-form makes the calibration of the model with ambiguity aversion a challenging task. A completely numerical calibration approach is very computationally demanding. Moreover, calibrated parameters of such a procedure can suffer from an inherent instability, because the objective function minimized by the calibration procedure depends itself on the solution (13) of another optimization problem. This feature implies a potentially non-smooth optimization problem, in which one numerical optimization is embedded in another one. We circumvent this problem by a simple analytical approximation, which exploits the fact that the ambiguity parameter η is typically a small number for application purposes.[17]

3.1.1 Yield curve asymptotics. We first assume that the gradient of the value function in Equation (13) can be approximated by a straightforward zero-order expansion:

$$V_Y(Y) = V_{0Y}(Y) + O(\sqrt{\eta}),$$ (28)

where V_{0Y} is the gradient of the zero-order value function in the economy with no ambiguity ($\eta = 0$). This gradient is known in closed-form.[18] Inserting the expansion (28) in Equation (15), the following first-order approximation for the market price of risk and ambiguity λ holds:[19]

$$\lambda = \sigma' + \sqrt{2\eta}\frac{\Xi'V_{0Y} + \sigma'}{\sqrt{(\Xi'V_{0Y} + \sigma')'(\Xi'V_{0Y} + \sigma')}} + o(\sqrt{\eta}). \tag{29}$$

Using this approximation, similar first-order expansions for the interest rate r and the change of drift ϕ_Y follow from Equations (16) and (18). By inserting these approximations in the fundamental pricing equation (19), we finally obtain a first-order expansion for the yield curve under ambiguity aversion. This expansion depends only on the zero-order value function gradient V_{0Y} and the zero-order yield curve ($\eta = 0$) in the completely affine model, which are both known in closed form. In this way, we can completely avoid a numerical solution of the optimization problem (11) when computing the yield curve under ambiguity aversion. The first-order approximation of the yield curve under ambiguity aversion is provided in Proposition 3.

Proposition 3. *Let $R(t,T) = -\frac{1}{T-t}\log P(t,T)$ be the zero coupon bond spot rate for maturity T. Then, the following first-order expansion holds:*

$$R(t,T) = R_0(t,T) - \sqrt{2\eta}\frac{P_1(t,T)}{(T-t)P_0(t,T)} + o(\sqrt{\eta}), \tag{30}$$

where R_0 and P_0 are the closed-form spot rate and zero coupon bond price in the model without ambiguity ($\eta = 0$). The first-order term $P_1(t,T)$ is given by:

$$P_1(t,T) = E_{\mathbb{Q}^0}\left[\int_t^T \exp\left(-\int_t^s (\alpha(u) - \sigma(u)\sigma(u)')du\right)\Psi_0(Y(s), s-t)ds\right], \tag{31}$$

where $\mathbb{Q}_0$ is the risk-neutral probability in the model without ambiguity, which is associated with the risk-neutral drift change $\phi_{0Y} = \Xi\sigma'$. The payoff function Ψ_0 in Equation (31) is defined for $t \leq s \leq T$ by:

$$\Psi_0(Y, s-t) = -\frac{(\Xi'V_{0Y} + \sigma')'[(s-t)\Xi'R_{0Y}(t,s) + \sigma']}{\sqrt{(\Xi'V_{0Y} + \sigma')'(\Xi'V_{0Y} + \sigma')}}P_0(t,s),\tag{32}$$

where $R_{0Y}(t,s)$ is the gradient of $R_0(t,s)$ with respect to the state variable Y.

Using Proposition 3, we easily obtain accurate first-order approximations for $R(t,T)$, by computing the expectation (31) with Monte Carlo methods. Since all terms defining Ψ_0 in Equation (32) are known in closed form for the completely affine model without ambiguity, this task can be accomplished quite efficiently.

3.1.1 Calibration results. When $\eta > 0$, we use the first-order yield curve approximation:

$$R(t,T) \approx R_0(t,T) - \sqrt{2\eta}\frac{P_1(t,T)}{(T-t)P_0(t,T)},\tag{33}$$

to calibrate our model to the U.S. Treasury yields data. When $\eta = 0$, we use the well-known closed-form expression $R_0(t,T)$ for the completely affine Longstaff and Schwartz (1992) term structure model. We calibrate the model to the unconditional mean and volatility of the one-month and the one-year yields, as well as to the first-order autocorrelations and the contemporaneous covariance of these yields. In order to provide additional information on the predictability structure in our interest rate data, we also match the Campbell and Shiller (1991) regression coefficients of two regressions of 10-year and 2-year yield changes on the past term structure slope, as measured with respect to the 6-month yield; see Subsection 3.2 for more details.

Table 1 summarizes the set of moment conditions used and the calibration results obtained for three distinct model settings. Setting I is based on an unconstrained calibration of the ambiguity parameter η. Setting II constrains the ambiguity parameter to an intermediate value between 0

and the unconstrained calibrated value obtained for Setting I. Setting III constrains the ambiguity parameter to 0, as in the completely affine model.

Insert Table 1 about here.

The optimal ambiguity parameter calibrated for Setting I. is $\eta = 0.0136$. The calibration results for the unconstrained model are rather good. The percentage calibration errors for the unconditional means and the contemporaneous covariance of the one-month and one-year yields, as well as for the Campbell and Shiller (1991) regression coefficients, are about 2%. The percentage errors for the unconditional volatilities and the first-order autocorrelation coefficients are of approximately 10%. Larger calibration errors are obtained in Settings II and III for the calibrated Campbell and Shiller (1991) regression coefficients and for the calibrated first-order autocorrelation coefficients. The poor performance in calibrating the Campbell and Shiller (1991) coefficients for settings in which $\eta \approx 0$ is consistent with the inability of completely affine models to match the empirical bond predictability patterns. However, the results in Table 1 highlight that a potentially small perturbation of the completely affine model can ameliorate substantially the empirical fit when ambiguity aversion is explicitly addressed in the definition of the model. To interpret and to better assess the size of the calibrated parameter $\eta = 0.0136$, we resort to model-detection probabilities, as proposed by Anderson, Hansen, and Sargent (2003). The goal of this exercise is to verify if the perturbed worst-case belief implied by our calibration is statistically sufficiently indistinguishable from the reference affine model. Table 2 reports detection error probabilities implied by our calibration for samples of increasing lengths.[20]

Insert Table 2 about here.

Anderson, Hansen, and Sargent (2003) advocate a detection-error probability of at least 10% for a reasonable entropy bound. As Table 2 shows, our calibrated model implies that the agent would

incorrectly reject the null that a given model holds with more than 10% probability, even with 50 years of available observations.

Using the calibrated parameters, we can also study the quality of the first-order yield curve approximation (30) for the ambiguity parameter $\eta = 0.0136$ implied by our unconstrained setting. We compute numerically the yield curve for several conditioning values of the state variables Y_1 and Y_2, by solving numerically Equations (13) and (19) at the calibrated model parameters. We then compare the yield curve with the one implied by the approximation (30). We find that our simple first-order approximation is quite accurate, even for longer maturities, with relative approximation errors typically below 10%. Figure 1 illustrates the quality of our yield curve approximation for three different yield curve states.

Insert Figure 1 about here.

The next section produces additional empirical evidence that our first-order yield curve approximation with ambiguity aversion improves the description of the salient features of U.S. Treasury yield data implied by completely affine models.

3.2 Empirical findings

We study the goodness-of-fit of the calibrated two-factor yield curve model with ambiguity aversion with respect to some well-known stylized facts of U.S. Treasury yields data.

3.2.1 Deviations from the expectations hypothesis: Campbell-Shiller Regressions. The expectations hypothesis of interest rates implies that bond returns are unpredictable. This hypothesis can be tested by a Campbell and Shiller (1991) time series regression of yield changes on the slope of the term structure:

$$R(t+m,t+n-m) - R(t,t+n) = \beta_0 + \beta_1 \frac{m}{n-m}[R(t,t+n) - R(t,t+m)] + \varepsilon_t, \qquad (34)$$

where n is the time to maturity of the zero bond and m is the length of the time period over which bond returns are measured. We fix $n = 1, 2, 3, 5, 7, 10$ years to maturity and m to be 6 months. Table 3 presents the coefficient estimates implied by our data set, together with those obtained for a long simulated sample of 5,000 observations from the calibrated models I, II, and III introduced in the last section. The coefficients for the 2-year and 10-year time to maturity, which have been used to calibrate the models, have been underlined.

Insert Table 3 about here.

The first panel of Table 3 emphasizes the way in which the expectations hypothesis is violated in the data. The estimated slope coefficients in these regressions are all significantly different from one: They are all negative and their absolute value increases with maturity. The last panel of Table 3 highlights the failure of the completely affine models to mimic the violations of the expectations hypothesis in the data: All estimated coefficients are positive and fall outside the confidence intervals estimated with our data set. The second panel of Table 3 shows that the model with an unconstrained calibrated ambiguity parameter can fit well the patterns of the estimated slope coefficients in the data: All estimated coefficients are negative and fall inside the confidence intervals estimated with our data set. The model with the constrained ambiguity parameter $\eta = 0.005$ also delivers negative estimated slope coefficients. However, in about half of the cases these estimates are outside the confidence intervals from the data.

These findings indicate that the model with the non-affine specification of the ambiguity premium can produce the predictability of bond excess returns consistent with the data. The distinct predictability patterns of the completely affine versus the ambiguity-averse version of the model are generated by the very different equilibrium excess returns implied by these two models. In both models, the market price of risk has a very simple structure, which is independent of the first state variable Y_1. However, in the model with ambiguity aversion, the ambiguity premium additionally influences excess returns in a nonlinear way. If follows that the ambiguity premium is not affine, depends on both state variables Y_1 and Y_2, and implies excess returns that can be both positive and

negative, in dependence of the realized state of the economy. Figure 2 plots the instantaneous risk and ambiguity premia $\Xi\lambda_R$ and $\Xi\lambda_A$ for Y_1 and Y_2, as a function of the relevant state (Y_1, Y_2) of the economy.

Insert Figure 2 about here.

The risk premium for the state Y_1 is always zero. The one for the state Y_2 is negative, independent of Y_1, and decreasing in Y_2. The negative risk premium for Y_2 is due to the negative correlation parameter ρ implied by the calibrated model. The ambiguity premia for Y_1 and Y_2 are both large and positive, and depend each nonlinearly on both state variables. The ambiguity premium for Y_1 is increasing in Y_1 and almost independent of Y_2. The ambiguity premium for Y_2 is increasing in Y_2 and decreasing in Y_1.

These features generate the more flexible structure of bond excess returns, which can accommodate the violations of the expectations hypothesis. To illustrate this point, Figure 3 plots the instantaneous expected excess return,

$$\mathbb{E}_t \left(\frac{dP(t, t+\tau)}{P(t, t+\tau)} \right) - r = \frac{P_Y(t, t+\tau)}{P(t, t+\tau)} \phi_Y,$$

of a 5-year maturity bond, as a function of the relevant state (Y_1, Y_2).

Insert Figure 3 about here.

The expected excess return is positive for a large set of possible states, it is increasing in Y_2 and decreasing in Y_1. Consistently with the empirical evidence, it can become negative for moderate values of Y_2 as the state Y_1 increases. Therefore, the model can generate predictability patterns driven by the two-dimensional system of state variables (Y_1, Y_2). In the two-factor model with ambiguity, this feature implies that the slope of the yield curve is a predictive factor for future bond returns.

3.2.2 Short interest rate dynamics. In the model with ambiguity, the market price of ambiguity influences in a direct way the level of the short rate by a non-affine function of the state variables. Therefore, it is interesting to study the properties of the short rate dynamics in this setting, and to compare them with those of the completely affine model with $\eta = 0$.

A useful tool to study these features is the pull function [see Conley, Hansen, Luttmer, and Scheinkman (1997)] which is a measure of the conditional speed of mean reversion for nonlinear diffusion processes.[21] The pull function $\mathcal{P}(r^\star)$ of the nonlinear diffusion process $r(t)$ is defined through the conditional probability that $r(t)$ reaches the value $r^\star + \varepsilon$ before $r^\star - \varepsilon$, if initialized at $r(0) = r^\star$, when ε is small. To first-order in ε, this probability is given by:

$$\frac{1}{2} + \varepsilon \frac{\mu_r(r^\star)}{2\sigma_r^2(r^\star)} + o(\varepsilon), \tag{35}$$

where μ_r and σ_r are the drift and diffusion functions of the short rate diffusion process, and the pull function is defined by:

$$\mathcal{P}(r^\star) = \frac{\mu_r(r^\star)}{2\sigma_r^2(r^\star)}. \tag{36}$$

We reproduce the pull function of the calibrated models with and without ambiguity by estimating it on a long sample of 25,000 simulated data from these models. The pull function is estimated using the semi-nonparametric method in Conley, Hansen, Luttmer, and Scheinkman (1997) for a flexible specification of the drift and the local volatility of the short rate; see Appendix B for details.[22]

Figure 4 presents the pull functions estimated for the short rate processes of the calibrated models with and without ambiguity (for $\eta = 0.0136$ and $\eta = 0$, respectively). For comparison, we also plot the empirical pull function and a 95% confidence interval around it, estimated using our data set.

Insert Figure 4 about here.

The pull function of the short rate process with ambiguity aversion is very similar to the one of the short rate in the calibrated two-factor completely affine model. Both pull functions highlight a pronounced nonlinearity of the mean reversion speed of the short rate in the calibrated models. The nonlinearity of the mean reversion obtained for the completely affine two-factor setting at the calibrated parameters is substantial. As shown in Buraschi and Jiltsov (2007), this feature could not have been generated, e.g., by calibrating a single-factor Cox, Ingersoll, and Ross (1985)-type short rate process, in which the pull function is explicitly given by:

$$\mathcal{P}(r^\star) = \frac{\lambda(\bar{r} - r^\star)}{2\sigma^2 r^\star}, \tag{37}$$

for some positive constants λ, σ, and $\bar{r}$. Over a good spectrum of short interest rate values, the pull functions of the calibrated models are contained in the 95% confidence interval around the empirical pull function. For short rate levels less than 4%, the model pull functions are outside the (large) confidence intervals. For these interest rate levels, however, it is possible that the point estimates and the estimated confidence intervals are not very reliable, due to the low fraction of short rate observations below 4% in our data set. Overall, these findings indicate that the yield curve model with the non affine specification of the ambiguity premium can reproduce the empirical failures of the expectations hypothesis without modifying in a substantial way the short rate nonlinear mean reversion features of the completely affine version of the model.

3.2.3 Yields volatility. The affine term structure literature documents the difficulties of low-dimensional affine factor models to match the dynamics of both the first and second moments of yields. Dai and Singleton (2003), for example, find that an $A_2(3)$ essentially affine model with one CIR and two Gaussian factors produces a conditional volatility that is approximately consistent with the data. However, this model fails largely in explaining the conditional first moments of bond yields. An $A_1(3)$ essentially affine model with two CIR volatility factors and one Gaussian factor matches even worse the volatility dynamics.

Is the time variation of the second moments in the calibrated model with ambiguity aversion roughly consistent with the empirical evidence? To investigate this issue we perform a simple exercise, as in Dai and Singleton (2003), and estimate a GARCH(1,1) model for the 5-year yield, using both our yield data and a large sample of 5,000 observations simulated at the calibrated model parameters.[23] The results of this exercise are summarized in Table 4.

Insert Table 4 about here.

In the calibrated model I., a moderate ambiguity ($\eta = 0.0136$) yields estimates of the GARCH parameters that are well within the 95%-confidence intervals around the GARCH point estimates in the data. The GARCH estimates for the calibrated completely affine model III ($\eta = 0$) are, instead, at odds with those estimated in the data. This model implies tight restrictions between bond excess returns and their volatility. The consequence of this feature is that it cannot match simultaneously the Campbell and Shiller (1991) regression coefficients and the volatility of interest rates. On top of the unsuccessful attempt to match the Campbell and Shiller (1991) regression coefficients in our calibration, this model fails with respect to the yield volatility dynamics. The model with ambiguity aversion breaks the link between bond excess returns and yield volatility. In this way, it can generate reasonable predictability and volatility patterns.

4. Conclusions

We extend completely affine continuous-time yield curve models to incorporate ambiguity (Knightian uncertainty) aversion, modeled by Multiple Priors Recursive Utility. This extension is parsimonious in the sense that it is parameterized by a single additional parameter: The degree of ambiguity in the model. In this setting, the excess returns on bonds reflect a premium for ambiguity, which is observationally distinct from the premium they pay for risk. Even in the simplest model, this ambiguity premium can be large. Moreover, it is also nonzero for state variables that

do not pay a premium for risk. We show that these features have non trivial implications for the model's ability to fit some stylized yield curve facts. We calibrate to U.S. Treasury yield data a simple two-factor model with ambiguity aversion. The model is able to reproduce the deviations from the expectations hypothesis documented in the literature, without modifying in a substantial way the nonlinear mean reversion dynamics of the short interest rate. In contrast to completely affine models, there is no apparent tradeoff between fitting the first and second moments of the yield curve. These findings suggest that a small degree of ambiguity can have large implications for explaining the yield curve stylized facts.

The insights gained from our analysis suggest interesting directions for future research. A key feature of our model is that state variables related to the production technology (e.g., via the drift of the production dynamics) can pay a time-varying premium for ambiguity, even if the technology dynamics itself is not ambiguous. Therefore, it would be interesting to model inflation as an ambiguous and heteroskedastic state variable influencing the drift of the cumulative return on capital. In this setting, the nominal term structure will be affected by a substantial and time-varying inflation premium, even if consumption growth itself could be difficult to forecast using its own history and is, if needed, homoskedastic. In this way, one could try to study the empirical term structure implications of ambiguity aversion in a structural model fully disciplined by macro data.[24]

Figure legends

Figure 1

First-order yield approximation error

The figure presents the exact yield curve and the first-order yield curve approximation in the calibrated model with ambiguity parameter $\eta = 0.0136$ for times to maturity τ given in years. The exact yield curves (continuous lines) are computed by solving numerically the differential equations (13) and (19) for the value function of the ambiguity-averse representative agent and for the yield curve under ambiguity aversion, respectively. The first-order yield curve approximations (dotted lines) are computed by Monte Carlo simulation using formula (31) in Proposition 3. The three panels in the figure plot the yield curves implied by three different realizations of the conditioning state variables Y_1 and Y_2.

Figure 2

Factor premia for risk and ambiguity

The figure plots the factor premia for risk and ambiguity of the state variables Y_1 and Y_2 (in the left and right panels, respectively). Factor premia for risk and ambiguity are plotted as functions of the state of the economy (Y_1, Y_2). The two panels on the top contain the factor premia for risk $\Xi\lambda_R$, those in the middle the factor premia for ambiguity $\Xi\lambda_A$. Finally, the total factor premia $\Xi\lambda$ are presented in the bottom panels.

Figure 3

Instantaneous bond expected excess returns

The figure plots the instantaneous expected excess return

$$\mathbb{E}_t \left(\frac{dP(t,t+\tau)}{P(t,t+\tau)} \right) - r = \frac{P_Y(t,t+\tau)}{P(t,t+\tau)} \phi_Y$$

of a bond with time to maturity $\tau = 5$ years, in dependence of the state variables Y_1 and Y_2, for the calibrated economy with ambiguity aversion ($\eta = 0.0136$).

Figure 4

Pull function of the short rate diffusion process

The figure plots the pull functions estimated using the approach of Conley, Hansen, Luttmer, and Scheinkman (1997). For the calibrated models with $\eta = 0$ and $\eta = 0.0136$, respectively, we simulate a time series of $25,000$ observations of the 1-month yield and then apply GMM to estimate the pull function from the simulated series. The dash-dotted line is the pull function of the completely affine model. The dash-boxed line is the one of the model with ambiguity aversion. In addition, we plot the pull function estimated by GMM for the time series of 1-month yields in our data set (solid line), together with the corresponding asymptotic 95% confidence intervals (dotted lines).

Appendix A: Proofs

These proofs were omitted from the main text.

Proof of Proposition 1.

Notice that our framework meets the regularity conditions required to apply the Saddle Point Theorem for infinite dimensional spaces [see Sion (1958) and Fan (1953)]. Therefore, we can alternatively characterize the value function $J(x,y)$ in (7) as:

$$J(x,y) = \inf_{h \in \mathcal{H}} \sup_{c,\pi} \quad \mathbb{E}^h \left[\int_0^\infty e^{-\delta t} \log(c(t)) dt \right] \quad . \tag{A1}$$

Let us first assume that the time horizon T is finite. According to the martingale formulation of the consumption-investment problem, to solve the first step of (A1), it is well known that optimality of c implies $c^*(t) = \exp(-\delta t)/(\xi_h(t)\psi)$, where the Lagrange multiplier ψ is solution of $\mathbb{E}^h \left[\int_0^T \xi_h(s) c^*(s) ds \right] = x$, i.e., $\psi = (1 - \exp(-\delta T))/\delta x$. $\xi_h(t)$ denotes the state price density for model P^h. This leads to:

$$c^*(t) = \delta \left(\frac{x e^{-\delta t}}{\xi_h(t)(1 - e^{-\delta T})} \right). \tag{A2}$$

Let:

$$J_h^T(x,y) = \mathbb{E}^h \left[\int_0^T e^{-\delta t} \log\left(c^*(t)\right) dt \right]. \tag{A3}$$

By virtue of (A2), one obtains:

$$J_h^T(x,y) = \frac{e^{-T\delta}(1 - e^{T\delta} + T\delta)}{\delta} + \log\left(\frac{\delta x}{1 - e^{-\delta T}}\right)\left(\frac{1 - e^{-\delta T}}{\delta}\right) \tag{A4}$$

$$+ \quad \mathbb{E}^h \left[\int_0^T e^{-\delta t} \int_0^t \left(r_h(s) + \frac{\theta_h(s)'\theta_h(s)}{2} \right) ds \, dt \right], \tag{A5}$$

where r_h and θ_h are the short rate and the market price for risk and ambiguity, respectively, for model P^h. In the infinite time horizon case, it follows that:

$$J_h(x,y) = \lim_{T \to \infty} J_h^T(x,y) = -\frac{1}{\delta} + \frac{\log(\delta x)}{\delta} + \mathbb{E}^h \left[\int_0^\infty e^{-\delta t} \int_0^t \left(r_h(s) + \frac{\theta_h(s)'\theta_h(s)}{2} \, ds \right) dt \right]. \tag{A6}$$

As a consequence of our inversion of the order of optimizations that leads to the value function (A1), we might consider a given Girsanov kernel h that satisfies (4) and the corresponding probability measure P^h. Within this model, we can infer from Cox, Ingersoll, and Ross (1985) the equilibrium interest rate process and excess return on financial assets. To this end, we recall the expression for the market price of risk and ambiguity of any admissible model P^h:

$$\theta_h(t) = \Sigma^{-1} \begin{pmatrix} \alpha - r \\ \beta - r\bar{1}_k \end{pmatrix} + h. \tag{A7}$$

We have:

$$r_h = \alpha - \sigma\sigma' + \sigma h, \tag{A8}$$

$$\beta_h = \alpha\bar{1}_k - \sigma(\sigma' - h)\bar{1}_k + \vartheta(\sigma' - h). \tag{A9}$$

Accordingly, the following equilibrium market price of risk also holds under P^h:

$$\lambda_h = \sigma' - h. \tag{A10}$$

It then follows that the following program gives the value function $J(x,y)$:

$$
\begin{aligned}
J(x,y) &= -\frac{1}{\delta} + \frac{\log(\delta x)}{\delta} + \inf_{h \in \mathcal{H}} \mathbb{E}^h \left[\int_0^\infty e^{-\delta t} \int_0^t \left(r_h(s) + \frac{\theta_h(s)'\theta_h(s)}{2} \right) ds\, dt \right] \\
&= -\frac{1}{\delta} + \frac{\log(\delta x)}{\delta} \\
&\quad + \inf_{h \in \mathcal{H}} \mathbb{E}^h \left[\int_0^\infty e^{-\delta t} \int_0^t \left(\alpha(Y(s)) - \frac{\sigma(Y(s))\sigma(Y(s))'}{2} + \sigma(Y(s)) \cdot h(s) \right) ds\, dt \right] \\
&= -\frac{1}{\delta} + \frac{\log(\delta x)}{\delta} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(A11)} \\
&\quad + \frac{1}{\delta} \inf_{h \in \mathcal{H}} \mathbb{E}^h \left[\int_0^\infty e^{-\delta t} \left(\alpha(Y(t)) - \frac{\sigma(Y(t))\sigma(Y(t))'}{2} + \sigma(Y(t)) \cdot h(t) \right) dt \right] \\
&= -\frac{1}{\delta} + \frac{\log(\delta x)}{\delta} + \frac{1}{\delta} V(y). \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(A12)}
\end{aligned}
$$

Dynamic programming implies the following necessary condition for optimality of h:

$$
\inf_{h \in \mathcal{H}} \left\{ V_Y' \left[\Lambda + \Xi h \right] + \frac{1}{2}\text{trace} \left[\Xi' V_{YY} \Xi \right] + \alpha - \frac{1}{2}\sigma\sigma' + \sigma \cdot h - \delta V \right\} = 0. \tag{A13}
$$

Due to the convexity in the control h of the functional that appears in curly brackets, the condition is also sufficient for optimality of h.[25] The complementary slackness condition that corresponds to the minimization (A13) implies:

$$
h^* = -\frac{1}{\psi} \left[\Xi' V_Y + \sigma' \right], \tag{A14}
$$

where:

$$
\psi = \frac{1}{\sqrt{2\eta}} \sqrt{\left(\Xi' V_Y + \sigma' \right)' \left(\Xi' V_Y + \sigma' \right)}. \tag{A15}
$$

Therefore, the process:

$$
h^* = -\sqrt{2\eta} \, \frac{\Xi' V_Y + \sigma'}{\sqrt{\left(\Xi' V_Y + \sigma' \right)' \left(\Xi' V_Y + \sigma' \right)}}, \tag{A16}
$$

constitutes an optimal feed-back control. We conclude that the value function of our model selection problem solves the nonlinear second-order Hamilton-Jacobi-Bellman PDE :

$$
V_Y' \Lambda + \frac{1}{2}\text{trace} \left[\Xi' V_{YY} \Xi \right] - \sqrt{2\eta} \sqrt{\left(\Xi' V_Y + \sigma' \right)' \left(\Xi' V_Y + \sigma' \right)} + \alpha - \frac{1}{2}\sigma\sigma' - \delta V = 0. \tag{A17}
$$

This concludes the proof.

$\square$

Proof of Corollary 1.

The equilibrium interest rate, premia on financial assets, and factor market prices of risk and ambiguity follow by substituting (A16) into the corresponding quantities that prevail under a generic admissible model P^h, i.e., (A8), (A9), and (A10). This concludes the proof.

$\square$

Proof of Proposition 3.

From Equation (29), we obtain directly the first order approximations for r and ϕ_Y:

$$r \;=\; \alpha - \sigma \left(\sigma' + \sqrt{2\eta} \frac{\Xi'V_{0Y} + \sigma'}{\sqrt{(\Xi'V_{0Y} + \sigma')'(\Xi'V_{0Y} + \sigma')}} \right) + o(\sqrt{\eta}), \tag{A18}$$

$$\phi_Y \;=\; \Xi \left(\sigma' + \sqrt{2\eta} \frac{\Xi'V_{0Y} + \sigma'}{\sqrt{(\Xi'V_{0Y} + \sigma')'(\Xi'V_{0Y} + \sigma')}} \right) + o(\sqrt{\eta}). \tag{A19}$$

We can insert these approximations in the fundamental pricing equation (19) for the case of the zero bond price. In this way, we can determine the first-order term in the expansion:

$$P(t,T) = P_0(t,T) + \sqrt{2\eta}P_1(t,T) + o(\sqrt{\eta}), \tag{A20}$$

by matching terms of same order in the fundamental pricing equation (19). Recalling that $P_0(t,T)$ solves the fundamental pricing equation for the economy without ambiguity ($\eta = 0$), we obtain the following partial differential equation for $P_1(t,T)$:

$$\frac{1}{2}\text{trace}\left(\Xi\,\Xi' \frac{\partial^2 P_1}{\partial Y \partial Y'} \right) + (\Lambda - \Xi\sigma')' \frac{\partial P_1}{\partial Y} - (\alpha - \sigma\sigma')P_1 + \frac{\partial P_1}{\partial t} = -\Psi_0, \tag{A21}$$

subject to the boundary condition $P_1(T,T) = 0$, where the payoff function Ψ_0 is defined in Equation (32). Now, we just remark that $r_0 = \alpha - \sigma\sigma'$ and $\phi_{0Y} = \Xi\sigma'$ are the short interest rate and the risk-neutral drift adjustment, respectively, of a Longstaff and Schwartz (1992) economy without ambiguity ($\eta = 0$). Therefore,

35

$P_1(t,T)$ can be interpreted as the price of a cash flow stream $\Psi_0(Y(s),s-t)$, $t \le s \le T$, in this economy. Using the Feynman-Kac formula, the expression for $P_1(t,T)$ in the proposition follows. This concludes the proof.

$\square$

Appendix B: GMM Estimation of the Pull Function

We estimate the pull function along the lines of the semi-nonparametric method in Conley, Hansen, Luttmer, and Scheinkman (1997). The estimator of the pull function at point $r^\star$ is:

$$\hat{\mathcal{P}}(r^\star) = \frac{\mu_r(r^\star;\hat{\theta})}{2\sigma_r^2(r^\star;\hat{\theta})} \, ,$$

where $\mu_r(r;\theta) = \sum_{j=-l}^{m} a_j r^j$ and $\sigma_r(r;\theta) = \exp\left[\sum_{i=0}^{p} c_j (\log r)^i\right]$ are flexible specifications for the drift and local volatility functions, and $\hat{\theta}$ is a GMM estimator of the parameter θ, containing the coefficients a_j, c_i, $j = -l,\ldots,m$ and $i = 0,\ldots,p$. We use $l = 1$, $m = 2$, $p = 2$, and the normalization $c_0 = 0$ for identification. This yields 6 parameters to be estimated. The GMM estimator $\hat{\theta}$ is based on the orthogonality condition:

$$E\left[\mu_r(r_t;\theta)\frac{d\varphi}{dr}(r_t) + \frac{1}{2}\sigma_r^2(r_t;\theta)\frac{d^2\varphi}{dr^2}(r_t)\right] = 0 \, ,$$

where r_t are discretely-sampled observations from process $r(t)$ and the 8×1 vector function φ is such that:

$$\frac{d\varphi}{dr}(r) = \frac{2}{\sigma_r^2(r;\theta)}\begin{bmatrix} H(r) \\ \sqrt{r}H(r) \end{bmatrix}, \tag{B1}$$

with $H(r) := \begin{bmatrix} \frac{1}{r} & 1 & r & r^2 \end{bmatrix}'$. For convenience, we have considered orthogonality conditions relying on the unconditional distribution of the short rate process. The first set of orthogonality conditions in Equation (B1) features an optimality property [see Conley, Hansen, Luttmer, and Scheinkman (1997)]. We use a two-step efficient GMM estimator, in which the optimal weighting matrix for the second step is obtained by

a Newey-West estimator with 5 lags. The asymptotic standard errors of $\hat{\mathcal{P}}(r^\star)$ are obtained (pointwise) from the GMM asymptotic variance-covariance matrix of $\hat{\theta}$ by applying the delta-method.

References

Aït-Sahalia, Y. 1996. Nonparametric Pricing of Interest Rate Derivative Securities. *Econometrica*, 64:527–560.

Aït-Sahalia, Y. 1999. Transition Densities for Interest Rate and Other Nonlinear Diffusions. *Journal of Finance*, 54:1361–1395.

Anderson, E. W., L. P. Hansen, and T. J. Sargent. 1998. Risk and Robustness in Equilibrium. Working Paper. University of Chicago.

Anderson, E. W., L. P. Hansen, and T. J. Sargent. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk and Model Detection. *Journal of the European Economic Association*, 1:68–123.

Buraschi, A., and A. Jiltsov. 2005. Inflation Risk Premia and the Expectations Hypothesis. *Journal of Financial Economics*, 75:429–490.

Buraschi, A., and A. Jiltsov. 2007. Habit Formation and Macroeconomic Models of the Term Structure of Interest Rates. *Journal of Finance*, 62:3009–3063.

Campbell, J. Y., and R. J. Shiller. 1991. Yield Spreads and Interest Rate Movements: A Birds Eye View. *Review of Economic Studies*, 58:495–514.

Cao, H., T. Wang, and H. Zhang. 2005. Model Uncertainty, Limited Market Participation and Asset Prices. *Review of Financial Studies*, 18:1219–1251.

Chen, Z., and L. G. Epstein. 2002. Ambiguity, Risk and Asset Returns in Continuous Time. *Econometrica*, 70:1403–1443.

Cheridito, P., D. Filipovic, and R. L. Kimmel. 2007. Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics*, 83:123–170.

Cochrane, J. H., and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review*, 95:138–160.

Conley, T. G., L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman. 1997. Short-Term Interest Rates as Subordinated Diffusions. *Review of Financial Studies*, 10:525–577.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407.

Dai, Q., and K. Singleton. 2002. Expectations Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure. *Journal of Financial Economics*, 63:415–441.

Dai, Q., and K. Singleton. 2003. Term Structure Modeling in Theory and Reality. *Review of Financial Studies*, 16:631–678.

Dow, J., and R. Werlang. 1992. Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio. *Econometrica*, 60:197–204.

Duffee, G. R. 2002. Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance*, 57:405-443.

Dybvig, J., and C. F. Huang. 1988. Nonnegative Wealth, Absence of Arbitrage, and Feasible Consumption Plans. *Review of Financial Studies*, 1:377–401.

Epstein, L. G., and J. Miao. 2003. A Two-Person Dynamic Equilibrium Under Ambiguity. *Journal of Economic Dynamics and Control*, 27:1253–1288.

Epstein, L. G., and M. Schneider. 2003. Recursive Multiple Priors. *Journal of Economic Theory*, 113:32–50.

Epstein, L. G., and T. Wang. 1994. Intertemporal Asset Pricing under Knightian Uncertainty. *Econometrica*, 62:283–322.

Fama, E. F., and R. R. Bliss. 1987. The Information in Long-Maturity Forward Rates. *American Economic Review*, 77:680–692.

Fan, K. 1953. Minimax Theorems. *Proceedings of the National Academy of Sciences of the United States of America*, 39:42–47.

Fleming, W. H., and H. M. Soner. 1993. *Controlled Markov Processes and Viscosity Solutions*. Springer, Berlin-New York.

Gagliardini, P., P. Porchia, and F. Trojani. 2004. Ambiguity Aversion and the Non Robustness of Some Affine Term Structures. Working Paper. University of St. Gallen.

Gilboa, I., and D. Schmeidler. 1989. Maxmin Expected Utility with Non-Unique Prior. *Journal of Mathematical Economics*, 18:141–153.

Kwon, H.-C. 1992. The Time Variant Term Premium of Interest Rates. Ph. D dissertation. Ohio State University Economics Department.

Leippold, M., F. Trojani, and P. Vanini. 2007. Learning and Asset Prices under Ambiguous Information. *Review of Financial Studies*, forthcoming.

Liu, J., J. Pan, and T. Wang. 2005. An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks. *Review of Financial Studies*, 18:131–164.

Longstaff, F. A., and E. S. Schwartz. 1992. Interest Rate Volatility and the Term Structure: a Two-Factor General Equilibrium Model. *Journal of Finance*, 47:1259–1282.

Maenhout, P. 2004. Robust Portfolio Rules and Asset Pricing. *Review of Financial Studies* 17:951–983.

McCulloch, H. 1990. US Term Structure Data, 1946–87. *Handbook of Monetary Economics*, I:672–715.

Sbuelz, A., and F. Trojani. 2002. Equilibrium Asset Pricing with Time Varying Pessimism. Working Paper. University of Lugano.

Sion, M. 1958. On General Minimax Theorems. *Pacific Journal of Mathematics*, 8:171-176.

Trojani, F., and P. Vanini. 2002. A Note on Robustness in Merton's Model of Intertemporal Consumption and Portfolio Choice. *Journal of Economic Dynamics and Control*, 26:423–435.

Trojani, F., and P. Vanini. 2004. Robustness and Ambiguity Aversion in General Equilibria. *The Review of Finance*, 2:279–324.

Uppal, R., and T. Wang. 2003. Model Misspecification and Under-Diversification. *Journal of Finance*, 59:2465–2486.

Wachter, J. 2006. A Consumption-Based Model of the Term Structure of Interest Rates. *Journal of Financial Economics*, 79:365–399.

Notes

[1]See, among others, Gilboa and Schmeidler (1989), Epstein and Wang (1994), Anderson, Hansen, and Sargent (1998, 2003), and Epstein and Schneider (2003).

[2]Epstein and Miao (2003) and Uppal and Wang (2003) show that ambiguity aversion can generate a home-bias and under-diversification in the optimal portfolios of the investors in a two-country economy. The equity premium and interest rate puzzles have been addressed, among others, in Chen and Epstein (2002), Maenhout (2004), Sbuelz and Trojani (2002), and Trojani and Vanini (2002). Leippold, Trojani, and Vanini (2007) show that the combination of learning and ambiguity aversion additionally can explain the excess volatility puzzle. The limited stock market participation generated by ambiguity aversion in the absence of market frictions has been early emphasized by Dow and Werlang (1992). Cao, Wang, and Zhang (2005) and Trojani and Vanini (2004) study ambiguity aversion and endogenous limited stock market participation in a static and a dynamic setting, respectively. Liu, Pan, and Wang (2005) show that the ambiguity about the jump probability in the return of the underlying asset can mimic the typical 'smirk' shape of options' implied volatilities.

[3]In their approach, the reference belief is interpreted as an approximate description of the unknown data-generating process. The discrepancy between models is measured by the relative entropy criterion.

[4]Aït-Sahalia (1996, 1999) and Conley, Hansen, Luttmer, and Scheinkman (1997) provide empirical evidence for a non-linearity in the drift of the short rate diffusion process.

[5]All coefficients to appear are assumed to be continuously differentiable functions of the state variables. Furthermore, we impose a uniform ellipticity condition on the matrix function $\Xi\Xi'$, where Ξ is the volatility matrix in Equation (3).

[6]We verify that in equilibrium drift and volatility coefficients of the financial assets return process depend on the state variable Y alone.

[7]Gagliardini, Porchia, and Trojani (2004) characterize the equilibrium with ambiguity aversion also in a finite time-horizon economy with an ambiguity averse representative agent that features a CRRA utility over terminal wealth. The logarithmic setting is convenient to emphasize the main intuition concerning ambiguity aversion and the term structure of interest rates, while preserving a higher tractability.

[8]See also Trojani and Vanini (2004), p. 289, for a more detailed discussion.

[9]An admissible trading strategy $\pi = [\omega \quad v]'$ is such that:

$$\int_0^t \left(|\omega(s)(\alpha(s) - r(s))| + |v(s)(\beta(s) - \overline{1}_k r(s))| + |\pi(s)'\Sigma(s)h(s)| + \left|\pi(s)'\Sigma(s)\Sigma(s)'\pi(s)\right|^2 \right) ds < \infty,$$

P-a.s. for every $t > 0$.

[10]See Appendix A for a formal justification of this step.

[11]Ambiguity aversion effects similar to second-order risk-aversion are precluded when the volatility of the production technology is stochastic, because the bound (4) implies a constant squared norm of the optimal drift contamination h^*.

[12]Note that the norm of vector λ_A is constant and equal to $\sqrt{2\eta}$. Therefore, λ_A is non-zero even when $\|\sigma(Y)\| \downarrow 0$.

[13]These equations can be solved in closed form only for particular model settings, e.g., by assuming a Gaussian dynamics for state variables Y or by introducing additional restrictions on the structure of the optimal drift distortion h^*. Gagliardini, Porchia, and Trojani (2004) provide a survey on such model settings.

[14]As shown in Gagliardini, Porchia, and Trojani (2004), ambiguity aversion generates an essentially affine term structure when all factors are Gaussian. Moreover, note that not all specifications

of ambiguity aversion in the literature are observationally distinct from second-order risk aversion behavior. Maenhout (2004), for example, models ambiguity aversion by a tilted robust control problem that is observationally equivalent to second-order risk aversion. For a setting with logarithmic utility, also the penalized control problem in Anderson, Hansen, and Sargent (1998, 2003), delivers a model of ambiguity aversion that is observationally equivalent to second-order risk-aversion behavior. Trojani and Vanini (2004) provide a detailed discussion of the observational distinctions of different models of ambiguity aversion.

[15]The detailed expressions for the solution are reported in Gagliardini, Porchia, and Trojani (2004).

[16]We thank Andrea Buraschi and Alexei Jiltsov for providing the dataset.

[17]The numerical value of η implied by all our calibrations is less than 0.0136. Moreover, note that our approximation method is conceptually different from an expansion with respect to the relevant state variables around some steady state level, like, for example, in the log-linearization approach.

[18]The explicit expression for $V_0(Y)$ is:

$$V_0(Y) = K_1 Y_1 + K_2 Y_2 + K_3\,,$$

with $K_1 = g_1/(m_1 + \delta)$, $K_2 = (g_2 - l^2)/(m_2 + \delta)$ and $K_3 = (a_1 K_1 + a_2 K_2)/\delta$.

[19]The higher order of the error in the approximation for λ, relative to the expansion (28), comes directly from the fact that the market price of risk and ambiguity depends on the gradient V_Y through a term that is pre-multiplied by $\sqrt{\eta}$.

[20]Anderson, Hansen, and Sargent (2003) provide a statistical model-detection tool to assess the adequacy of a given entropy bound. This tool is a measure of the similarity between the reference

belief and the worst-case belief selected by the agent. Consider the process for the log-likelihood ratio between model P^{h^*} and P:

$$\varsigma_1(t) = \log \frac{dP^{h^*}}{dP} = -\int_0^t h^*(s)' dZ(s) - \frac{1}{2}\int_0^t \|h(s)\|^2 ds.$$

Similarly, the log-likelihood ratio between model P and P^{h^*} is given by:

$$\varsigma_2(t) = \log \frac{dP}{dP^{h^*}} = \int_0^t h^*(s)' dZ(s) + \frac{1}{2}\int_0^t \|h(s)\|^2 ds = -\varsigma_1(t).$$

If model P holds, the agent will erroneously reject it based on a sample of observations of length $[0,T]$ when $\varsigma_1(T) > 0$. Conversely, if model P^{h^*} holds and $\varsigma_2(T) > 0$ (or, equivalently, $\varsigma_1(T) < 0$), the agent will erroneously reject it. The detection error probability is the conditional probability of the occurrence of these events, when equal priors are assigned to competing models:

$$p_T(\eta) = \frac{1}{2}P\left[\xi_1(T) > 0 \mid \mathcal{F}_0\right] + \frac{1}{2}P^{h^*}\left[\xi_1(T) < 0 \mid \mathcal{F}_0\right].$$

[21]A different approach to study the properties of nonlinear diffusion processes has been proposed by Aït-Sahalia (1996, 1999).

[22]The dynamics of the short interest rate implied by our two-factor model is not autonomous. Like in indirect inference estimation, the semi-nonparametric model defining the pull function estimates can be interpreted as a flexible auxiliary model for the short rate diffusion process.

[23]Note that the 5-year yield has not been used to calibrate our model to the yield curve data.

[24]We thank an anonymous referee for having pointed out this possible application of our framework.

[25]See Fleming and Soner (1993), Theorem 3.1.

Table 1

Calibrated unconditional moments and percentage calibration errors

	Maturities	Calibration Error (%)					
		1		2		3	
Mean of $R(t,t+\tau)$	$\tau = $ 1M, 1Y	-1.91	-2.87	-1.48	-1.95	-1.24	-2.02
Volatility of $R(t,t+\tau)$	$\tau = $ 1M, 1Y	-10.41	-9.15	-12.40	-13.80	-9.40	-5.30
Cov. of $R(t,t+\tau_1)$, $R(t,t+\tau_2)$	$\tau_1 = $ M, $\tau_2 = $ 1Y	1.93		2.38		3.30	
Autocorr. of $R(t,t+\tau)$	$\tau = $ 1M, 1Y	10.32	7.31	14.30	10.23	26.30	17.30
C. S. coeff. β_1	$n = $ 2Y, 10Y; $m = $ 6M	-1.14	-0.90	-76.72	-7.77	-129.77	153.72

The table displays results of the three different calibrations. The first and the second columns list the set of unconditional moments and maturities used in the calibration. These are the average 1-month and 1-year yields to maturity, the sample volatility of the 1-month and 1-year yields to maturity, the sample covariance between 1-month and 1-year yields to maturity, the first-order autocorrelation of the 1-month and 1-year yields to maturity and, finally, the Campbell and Shiller (1991) coefficients for the regression of 10-year and 2-year yield changes on the slope of the term structure, as measured by the difference between the 10-year and 2-year yields, respectively, and the 6-months yield:

$$R(t+m,t+n-m) - R(t,t+n) = \beta_0 + \beta_1 \frac{m}{n-m}[R(t,t+n) - R(t,t+m)] + \varepsilon_t .$$

The remaining columns present calibration percentage errors for each calibrated unconditional moment and Campbell and Shiller coefficient across the different model settings. Column 1 presents results for the calibration with an unconstrained η parameter, where at the optimum $\eta = 0.0136$. Column 2 presents results for the calibration with a constrained parameter $\eta = 0.005$. Column 3 presents results for the calibration with constrained parameter $\eta = 0$.

Table 2

Detection-error probabilities

	T=120 months	*T*=300 months	*T*=600 months
$p_T(\eta)$	0.1925	0.1704	0.1145

Detection-error probabilities for the calibrated model with ambiguity aversion ($\eta = 0.0136$), based on sample sizes of $T = 120$ months, $T = 300$ months, and $T = 600$ months, respectively. Detection-error probabilities have been estimated by Monte-Carlo methods with 5,000 simulated paths of the log-likelihood ratio and a time-step of $1/120$ of a month.

Table 3
Campbell and Shiller (1991) regression coefficients

Maturity	$n=1Y$	$n=2Y$	$n=3Y$	$n=5Y$	$n=7Y$	$n=10Y$
		Data				
β_1	-0.5784	-0.9548	-1.2386	-1.7234	-2.1350	-2.6219
95% CI	(-1.3241, 0.1672)	(-1.8545,-0.0552)	(-2.2630,-0.2142)	(-2.9644,-0.4824)	(-3.5288,-0.7412)	(-4.3212,-0.9226)
		Model I ($\eta=0.0136$)				
β_1	-0.4355	-0.94391	-1.1146	-1.5624	-1.7205	-2.6454
		Model II ($\eta=0.005$)				
β_1	-0.0431	-0.0860	-0.1172	-0.2142	-0.3062	-2.8256
		Model III ($\eta=0$)				
β_1	0.142	0.2842	0.4022	0.7106	0.9926	1.4084

The table presents the results of the Campbell and Shiller (1991) regressions:

$$R(t+m,t+n-m) - R(t,t+n) = \beta_0 + \beta_1 \frac{m}{n-m}[R(t,t+n) - R(t,t+m)] + \varepsilon_t,$$

for the three different calibrations we performed. The maturities $n = 1,2,3,5,7,10$ are in years. The Campbell and Shiller (1991) coefficients already used in the model calibrations are those for $n = 2,10$ and are underlined. The investment horizon m is 6 months in all regressions. In the first panel, we report the estimated coefficients for our sample of U.S. Treasury yields (first row), together with the corresponding 95% confidence intervals in parentheses (second row). All confidence intervals are computed using the Newey-West variance-covariance matrix estimator with 8 lags. In the second panel, we present the estimated coefficients for a long simulated sample from the unconstrained calibrated model with η = 0.0136. The third panel presents the estimated coefficients for a long simulated sample from the constrained calibrated model with η = 0.005. The fourth panel presents the estimated coefficients for a long simulated sample from the constrained completely affine Longstaff and Schwartz (1992) model without ambiguity (η = 0).

48

Table 4

GARCH(1,1) parameters for the 5-year yield

	α	β
Data	0.0844	0.90231
	(0.0421, 0.1267)	(0.8328, 0.9718)
Model I. ($\eta = 0.0136$)	0.06281	0.92361
Model III. ($\eta = 0$)	0.25624	0.41376

The table presents the point estimates for the GARCH(1,1) part of the following AR(1)-GARCH(1,1) model:

$$
\begin{aligned}
R(t, t+\tau) &= a_1 + a_2 R(t-1, t-1+\tau) + \sigma_t \varepsilon_t \\
\sigma_t^2 &= c + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 .
\end{aligned}
$$

The first row presents the point estimates obtained using the time series of $\tau = 5$-year maturity yields in our data set (95% confidence intervals are given in parentheses). In the second row, the GARCH parameters are estimated with a sample of 5,000 observations for the $\tau = 5$-year maturity yield simulated from the calibrated model with ambiguity aversion ($\eta = 0.0136$). In the third row, the GARCH parameters are estimated with a sample of 5,000 observations simulated from the calibrated Longstaff and Schwartz (1992) model ($\eta = 0$).
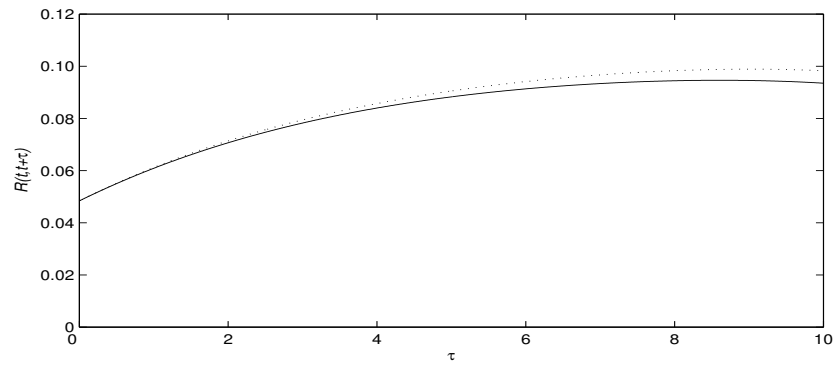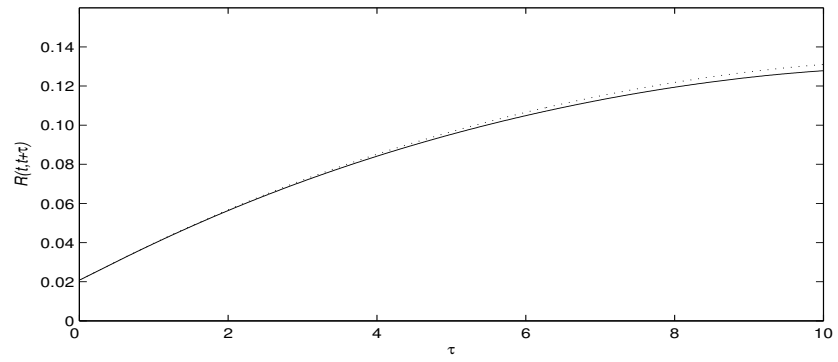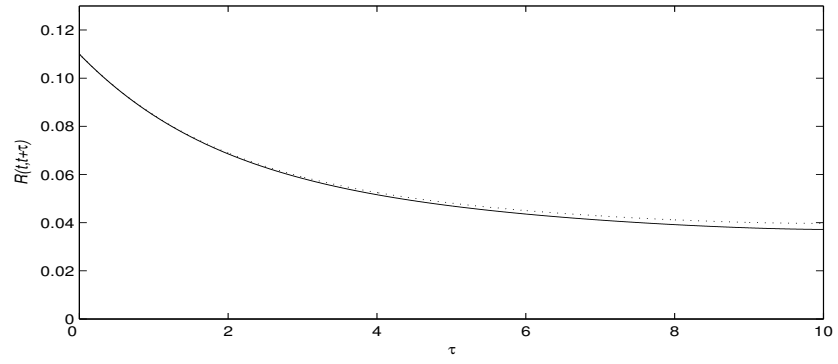
Figure 1
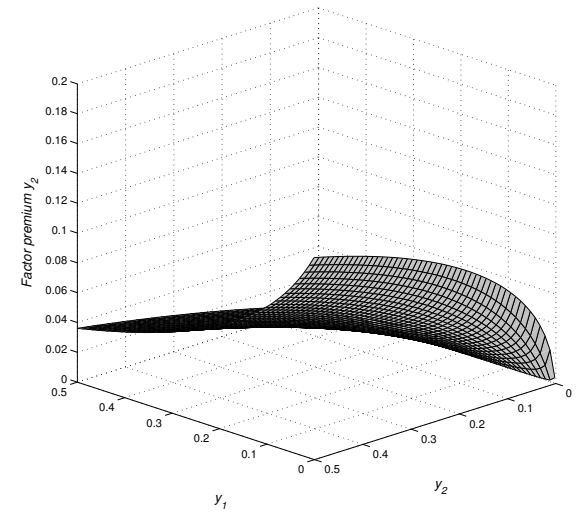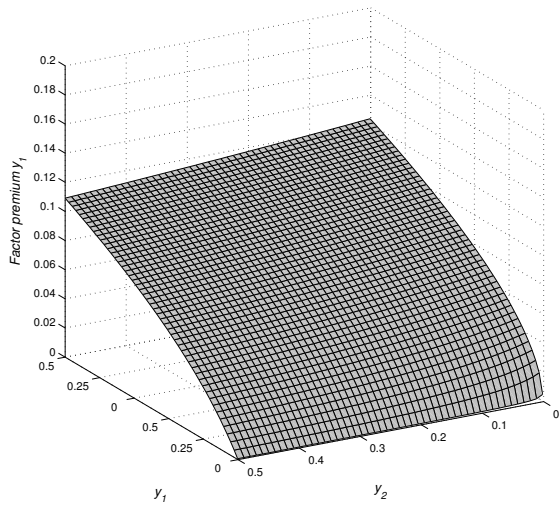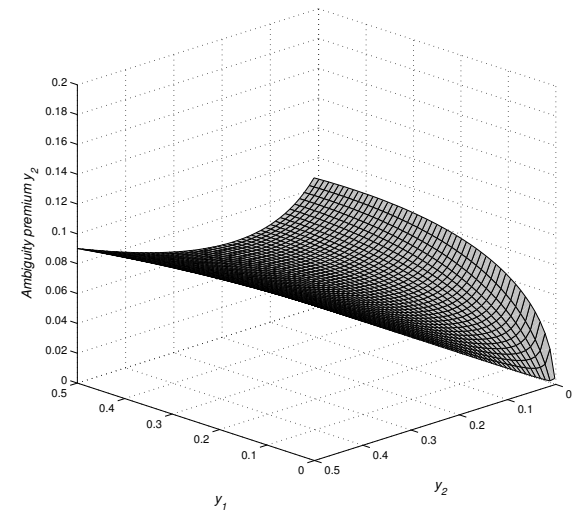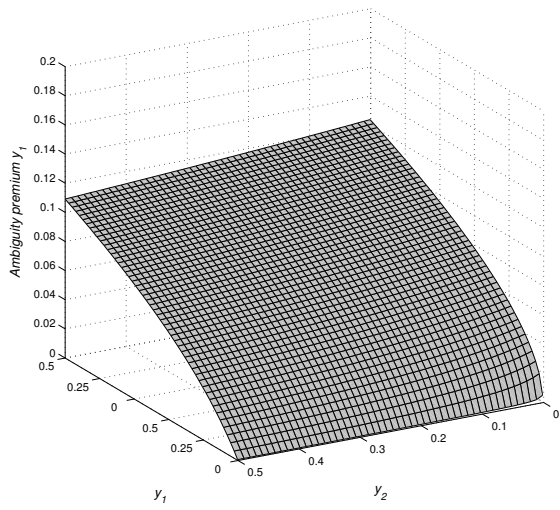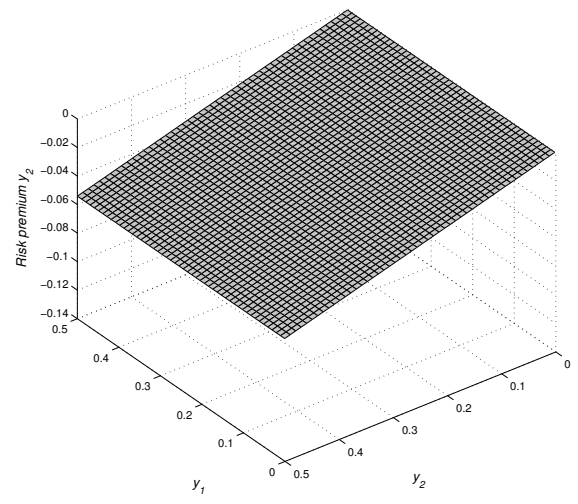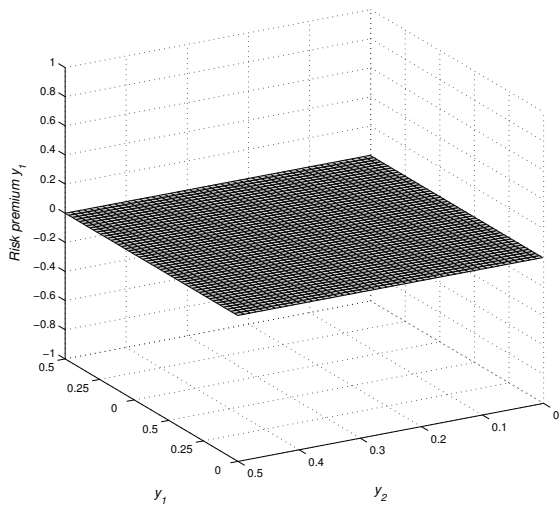
Figure 2

Figure 3

Figure 4