Research paper

Computing equilibrium bond prices in the Vayanos-Vila model☆

Fumio Hayashi

National Graduate Institute for Policy Studies 7-22-1 Roppongi, Minato-ku, Tokyo, 106-0032, Japan

A B S T R A C T

We develop tools for computing equilibrium bond prices for the discrete-time version of the Vayanos–Vila (2009) model. With the maturity structure included in pricing factors, factor loadings for equilibrium bond yields depends critically on parameters describing maturity structure dynamics and other model parameters. An illustrative example shows that the effect on the yield curve of a supply shock originating in a given maturity, although hump-shaped around the originating maturity, is to change yields broadly across all maturities.

1. Introduction and summary

Governments often attempt to influence the yield curve through open-market operations. This paper theoretically studies the effect of changes in the maturity structure engineered by such policy actions. It does so by developing tools for computing equilibrium bond prices for the explicit model of bond markets of Vayanos and Vila (2009).

The Vayanos–Vila model has two types of investors, risk-averse arbitrageurs and a set of maturity-specific "preferred-habitat" investors who differ from each other in the preferred maturity of bonds they hold. Equilibrium yields are affine functions of pricing factors including supply factors affecting the net supply of bonds by preferred-habitat investors. If the model is specialized so that the net supply is price-inelastic, then the preferred habitat investors as a whole can be interpreted as the government exogenously supplying bonds of different maturities, with the supply factors driving the entire maturity structure.

This specialized model has been put to use in recent empirical studies. The dynamics of the maturity structure in Greenwood and Vayanos (2014) and Greenwood et al. (2015) is very restrictive: only one supply factor drives the whole profile of the maturity structure, even though (as remarked in Greenwood and Vayanos (2014, p. 12)) multiple factors might

be needed to describe supply. D'Amico and King (2013), in their analysis of the recent Federal Reserve's large-scale purchase of Treasury securities, considered the supply effect of government bonds of a wide variety of maturities. To compute the factor loadings for equilibrium yields, they resort to an approximation suggested by Vayanos and Vila (2009). The paper by Hamilton and Wu (2012) is another empirical study of the supply effect motivated by the Vayanos–Vila model, but their regressions linking bond supply to bond prices are not derived from the model.

The empirical applications are constrained thus because computing equilibrium bond prices in the Vayanos–Vila model with more-than-several pricing factors has been viewed as intractable. The difficulty lies in the system of equations for the affine factor loadings. Unlike in the standard ATSM (affine term structure model) in modern finance, it is a non-linear system and cannot be solved recursively.

The contribution of this paper is to provide tools for overcoming this computational difficulty, for the discrete-time version of the Vayanos–Vila model. There is no need to restrict the maturity structure dynamics except that it is an unrestricted VAR (vector autoregression). The key observation is that the nonlinear factor loadings equations are QVEs (quadratic vector equations). We show that the solution can be obtained via a system of ODEs (ordinary differential equations) in a *single* variable. We verify that, for a set of model parameters we examine, the solution thus obtained can be located quickly by a fixed-point algorithm, a finding which should prove useful in a structural estimation of the model.

Including a large number of pricing factors in the affine yield equations, made feasible by the tools just described, is consistent with the well-known empirical fact that only up to three factors are enough to describe the yield curve. This is because the affine factor loadings, being a solution to the model, are not free parameters. The model's parameters are the VAR parameters describing maturity structure dynamics and a single risk-aversion parameter for the arbitrageurs. Therefore, there is a set of tight cross-equation restrictions, controlled by the single parameter, between the VAR and equations linking the pricing factors to the yield curve. In this sense, the model is a one-factor model.

As an illustration, we compute equilibrium bond prices for the case of maximal number of supply factors. That is, the vector of supply factors is the maturity structure itself. This enables us to examine two aspects of the bond market that have eluded rigorous treatment. First, the maturity structure dynamics can accommodate what we call the "legacy" property of bonds, which is that an increase in the supply of n-period bonds *necessarily* leads to an increase, in the next period, of the same amount in the supply of $(n-1)$-period bonds, unless there is an offsetting market operation. Second, we can give a rigorous formulation of a "local" supply shock. Shocks to the maturity structure, one for each maturity, are local in that they are uncorrelated across maturities. Neither aspect would be difficult to analyze if the maturity structure were driven by a small number of supply factors.

Our computation shows that the response of bond yields to a local supply shock originating at a given maturity occurs broadly across all maturities but is hump-shaped, with the maximum response at around the originating maturity. Consider, for example, a shock to the supply of the 10-year bond. One may wonder why the yield on the 20-year bond does not go up more than the yield of the 10-year bond because supply affects risk premia which should manifest themselves more for bonds with longer duration. The key intuition is that the shock will affect the 20-year bond only for 10 years because the shock will cease to exist after that. This is a useful result, especially given the empirical evidence in, e.g., D'Amico and King (2013) that supply effects have their maximum effects on yields at the maturities where they originate. We will also show that the hump-shaped response of the yield curve remains when supply shocks are correlated across maturities.

The rest of the paper is organized as follows. To place the Vayanos–Vila model in a proper perspective, we provide a brief review of the relevant literature in Section 2. Section 3, in order to set a stage for the development of our computational tools in Section 4, presents the Vayanos–Vila model in discrete time. Section 5 examines the effect of supply shocks for the illustrative example. Section 6 states the conclusion and an agenda for future research.

2. Modeling the supply effect on the yield curve

This section briefly reviews the literature on the supply effect of government bonds on the yield curve. This topic has attracted renewed interests in recent years with the advent of the U.S. Federal Reserve's LSAP (Large-Scale Asset Purchases) and the Bank of Japan's QQE (quantitative and qualitative easing), both of which involve a very large amount of purchases of long-term government bonds with the stated intention of lowering long-term interest rates.

The topic has been studied by prominent researchers for decades. Early contributions include Modigliani and Sutch (1966), whose preferred-habitat model assumes that the bond market is segmented by the maturity so that the interest rate for a given maturity is determined by the supply and demand for bonds of the maturity. In Tobin (1969) portfolio-balance model, the market segmentation is not as extreme because the demand for bonds of a given maturity can depend on the interest rates for other maturities, but it still is the case that the equilibrium term structure of interest rates is determined by the relative supply of bonds of different maturities. This view of the government bond market is widely held by central bank policymakers. For example, in his Jackson Hole speech in August 2010, the then Fed chairman Ben Bernanke stated:

> "… the Fed's strategy relies on the presumption that different financial assets are not perfect substitutes in investors portfolios, so that changes in the net supply of an asset available to investors affect its yield and those of broadly similar assets." (Bernanke, 2010)

In sharp contrast, the academic literature has departed from this view, probably because its underlying models have no micro-foundations. Current models are built on the no-arbitrage condition, which requires that the interest rates differ from each other only because of their different relation to a common factor called the SDF (stochastic discounting factor). In one class of models, the SDF is driven by a vector of pricing factors that follow a specific stochastic process, which leads to the celebrated ATSM (affine term structure model) in the finance literature that the equilibrium yields are affine in the pricing factors. It is not clear, however, why bond supply matters in the no-arbitrage framework. One could proceed with the ATSM, as, for example, in Li and Wei (2013), by including bond supply in the pricing factors, but it leaves unanswered the question of why bond supply is a pricing factor.

The model of Vayanos and Vila (2009) is attractive for a number of reasons. First, it is an explicit model of the bond market with micro-foundations that formalizes the preferred-habitat and portfolio balance effects. Second, as a result, it is very clear about why bond supply could matter. Third, it is an ATSM, so the implied yield equations are directly comparable to those in the ATSM literature.

The resemblance of the Vayanos–Vila model to the ATSM, however, is superficial because of its complexity in the computation of the affine coefficients in the yield equations. The computed example by Vayanos and Vila (2009), in its Section 5.5 is for a very limited case of only two maturities. The computational complexity has limited the empirical applicability of the Vayanos–Vila model. The present paper removes the limitation by providing computation tools that can handle a large number of maturities.

3. The specialized Vayanos–Vila model in discrete time

The bond pricing model to be considered in this paper is the discrete-time version of Vayanos and Vila (2009) with the added restriction that the preferred-habitat investors are replaced by an entity interpretable as the government supplying bonds inelastically. This section presents the model and a system of equations for equilibrium bond prices. The equation system will be the object of analysis in the next section.

3.1. The discrete-time bond pricing model

Denote by $P_t^{(n)}$ the price of zero-coupon government bonds with maturity n in period t. The (continuously compounded) yield to maturity $y_t^{(n)}$ is related to the bond price $P_t^{(n)}$ as in

$$y_t^{(n)} = -\frac{1}{n} \log P_t^{(n)}. \tag{3.1}$$

The nominal share of the n-period bond supplied, defined as the value of n-period bonds divided by the total value of bonds of all maturities supplied, is denoted by $s_t^{(n)}$. The longest maturity of bonds supplied is N. The shares add up to one: $\sum_{n=1}^{N} s_t^{(n)} = 1$. The $N-1$ dimensional vector $(s_t^{(2)}, s_t^{(3)}, \ldots, s_t^{(N)})$ is the *maturity structure* of bonds supplied.

Those bonds are traded by arbitrageurs. If $z_t^{(n)}$ is the nominal share of their n-period bond holdings, the holding-period return on their bond portfolio is

$$R_{t+1} \equiv \sum_{n=1}^{N} \frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}} z_t^{(n)}. \tag{3.2}$$

The decision problem of arbitrageurs is to maximize the risk-adjusted portfolio return subject to the adding-up constraint:[1]

$$\max_{\{z_t^{(n)}\}_{n=1}^{N}} \left[E_t(R_{t+1}) - \frac{\gamma}{2} \text{Var}_t(R_{t+1}) \right] \qquad \text{subject to} \qquad \sum_{n=1}^{N} z_t^{(n)} = 1. \tag{3.3}$$

Here, γ (≥ 0) is a risk-aversion parameter. The FOCs (first-order conditions), derived in Appendix A for the sake of completeness, are

$$E_t \left[\underbrace{\frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}}}_{\text{holding-period return}} \right] - \underbrace{\frac{1 - P_t^{(1)}}{P_t^{(1)}}}_{\text{short rate}} = \gamma \frac{1}{2} \frac{\partial \text{Var}_t(R_{t+1})}{\partial z_t^{(n)}}, \quad n = 2, 3, \ldots, N, \tag{3.4}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{risk premium on } n\text{-period bonds}}$$

[1] We follow the discrete-time formulation by Hamilton and Wu (2012) of the Vayanos-Vila model to take the R_{t+1} in the objective function to be the portfolio return rather than the increase in wealth. In the latter case, $z_t^{(n)}$ would be the level of nominal investment in bonds of maturity n. We are thus normalizing R_{t+1} and $z_t^{(n)}$ to the initial wealth. We do this because the model's calibration later in the paper is more straightforward with this normalization.

which state that, for arbitrageurs to hold a given amount of n-period bonds, the risk premium on those bonds must be enough to compensate for an increase in the volatility of the portfolio return R_{t+1} due to the inclusion of those n-period bonds in the portfolio.

The bond market equilibrium is that $z_t^{(n)} = s_t^{(n)}$ for $n = 1, 2, \ldots, N$. Because of the adding-up constraint $\sum_{n=1}^{N} z_t^{(n)} = 1$, only $N - 1$ of these equilibrium conditions are independent. The choice (to be mentioned right below) of factors for bond pricing implies that the price of one-period bonds, $P_t^{(1)}$, is determined outside the model. Given $P_t^{(1)}$, the $N - 1$ equilibrium conditions determine $N - 1$ bond prices $(P_t^{(2)}, \ldots, P_t^{(N)})$.

Equilibrium bond prices are assumed to depend on a vector, $\mathbf{f}_t$, of pricing factors whose dynamics is given by a Gaussian VAR(1):

$$\mathbf{f}_{t+1} = \mathbf{c} + \mathbf{\Phi}\,\mathbf{f}_t + \boldsymbol{\varepsilon}_{t+1}, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega}). \tag{3.5}$$

The shortest yield $y_t^{(1)}$ is included in $\mathbf{f}_t$, which means that the short rate, and hence the price of one-period bonds $P_t^{(1)}$, are given to the model. The remaining factors are denoted by the K-dimensional vector $\boldsymbol{\beta}_t$. Thus

$$\underset{((1+K)\times 1)}{\mathbf{f}_t} \equiv \begin{bmatrix} y_t^{(1)} \\ \boldsymbol{\beta}_t \\ {\scriptstyle (K\times 1)} \end{bmatrix}. \tag{3.6}$$

We will call the factors included in $\boldsymbol{\beta}_t$ the *supply factors* because they span the maturity structure as in

$$\underbrace{\begin{bmatrix} s_t^{(2)} \\ \vdots \\ s_t^{(N)} \end{bmatrix}}_{((N-1)\times 1)} = \underset{((N-1)\times K)}{\mathbf{C}} \underset{(K\times 1)}{\boldsymbol{\beta}_t} = \underset{((N-1)\times(1+K))}{\mathbf{S}} \underset{((1+K)\times 1)}{\mathbf{f}_t}, \tag{3.7}$$

where

$$\underset{((N-1)\times(1+K))}{\mathbf{S}} \equiv \begin{bmatrix} \underset{((N-1)\times 1)}{\mathbf{0}} & \underset{((N-1)\times K)}{\mathbf{C}} \end{bmatrix}. \tag{3.8}$$

3.2. The factor loadings equations

Vayanos and Vila (2009) have shown for the continuous-time case that equilibrium bond prices are exponential affine in $\mathbf{f}_t$:

$$\log P_t^{(n)} = \bar{a}_n + \bar{\mathbf{b}}_n' \mathbf{f}_t, \quad \text{so by (3.1)}, \quad y_t^{(n)} = a_n + \mathbf{b}_n' \mathbf{f}_t \quad \text{where} \quad a_n \equiv -\frac{\bar{a}_n}{n}, \quad \mathbf{b}_n \equiv -\frac{\bar{\mathbf{b}}_n}{n}, \\ n = 1, 2, \ldots, N. \tag{3.9}$$

They also derived a system of equations that have to be satisfied by the affine factor loadings (i.e., the coefficients $(\bar{a}_n, \bar{\mathbf{b}}_n)$ in the above affine function).

For the discrete-time model at hand, the affine functional form can be derived as an approximation by log-linearizing. The factor loadings equations, derived in Appendix A for completeness's sake, are

$$\underset{(1\times F)}{\bar{\mathbf{b}}_n'} = \underset{(1\times F)}{\bar{\mathbf{b}}_{n-1}'} \underset{(F\times F)}{\mathbf{\Phi}} - \gamma \underset{(1\times F)}{\bar{\mathbf{b}}_{n-1}'} \underset{(F\times F)}{\mathbf{\Omega}} \underbrace{\begin{bmatrix} \underset{(F\times 1)}{\bar{\mathbf{b}}_1} & \underset{(F\times 1)}{\bar{\mathbf{b}}_2} & \cdots & \underset{(F\times 1)}{\bar{\mathbf{b}}_{N-1}} \end{bmatrix}}_{(F\times(N-1))} \underset{((N-1)\times F)}{\mathbf{S}} + \underset{(1\times F)}{\bar{\mathbf{b}}_1'}, \tag{3.10}$$

$$\bar{a}_n = \bar{a}_{n-1} + \underset{(1\times F)}{\bar{\mathbf{b}}_{n-1}'} \underset{(F\times 1)}{\mathbf{c}} + \frac{1}{2} \underset{(1\times F)}{\bar{\mathbf{b}}_{n-1}'} \underset{(F\times F)}{\mathbf{\Omega}} \underset{(F\times 1)}{\bar{\mathbf{b}}_{n-1}} + \bar{a}_1, \quad n = 2, 3, \ldots, N, \tag{3.11}$$

where F is the number of factors (which equals $1 + K$ where K is the number of supply factors). Given an initial condition about $\bar{\mathbf{b}}_1$, Eq. (3.10) determines $(\bar{\mathbf{b}}_2, \ldots, \bar{\mathbf{b}}_N)$. Given an initial condition about $\bar{a}_1$ and given $(\bar{\mathbf{b}}_1, \ldots, \bar{\mathbf{b}}_N)$ thus determined, the recursion (3.11) generates $\bar{a}_n$, $n = 2, 3, \ldots, N$. Since $y_t^{(1)} = -\log P_t^{(1)}$ and since $y_t^{(1)}$ is the first element of $\mathbf{f}_t$, it follows from (3.9) for $n = 1$ that

$$\bar{a}_1 = 0, \quad \bar{\mathbf{b}}_1 = (-1, 0, \ldots, 0)', \tag{3.12}$$

which provides the initial conditions.

3.3. Yields and risk premia

For later use in Section 5, we point to a formula showing how yields are connected to risk premia. Let $\mathrm{rp}_t^{(n)}$ ($n = 2, 3, \ldots, N$) stand for the risk premium on n-period bonds ($\mathrm{rp}_t^{(1)} = 0$ by definition). The usual forward iteration argument

(included in Appendix B) then yields:

$$y_t^{(n)} = \underbrace{\frac{1}{n}\sum_{k=0}^{n-1}\mathrm{E}_t\left(y_{t+k}^{(1)}\right)}_{\text{expectations hypothesis benchmark}} + \underbrace{\frac{1}{n}\sum_{k=0}^{n-2}\mathrm{E}_t\left(\mathrm{rp}_{t+k}^{(n-k)}\right)}_{\text{"term premium"}} + \frac{1}{n}A_n, \quad n = 2, 3, \ldots, N, \tag{3.13}$$

where $A_n \equiv -\frac{1}{2}\sum_{k=0}^{n-2}\bar{\mathbf{b}}_{n-k-1}'\mathbf{\Omega}\bar{\mathbf{b}}_{n-k-1}$ is the Jensen inequality term. This formula shows that the yield on an n-period bond has two time-varying components. One is the expectations hypothesis benchmark, namely the average over the bond's life of the current and future short rates. The other component, which is sometimes called the term premium, can be written as the average of the *longitudinal* profile of the risk premium, $\left(\mathrm{rp}_t^{(n)}, \mathrm{rp}_{t+1}^{(n-1)}, \ldots, \mathrm{rp}_{t+n-2}^{(2)}\right)$, on the n-period bond as it approaches maturity.

4. Solving the quadratic vector equation (3.10)

4.1. The QVE

Unlike in the standard textbook ATSM (affine term structure model) and in contrast to (3.11), Eq. (3.10) is not a recursion thanks to the quadratic term (the second term on the right-hand side of the equation) that involves the whole array $(\bar{\mathbf{b}}_1, \ldots, \bar{\mathbf{b}}_{N-1})$. To make it clear that it is a QVE (quadratic vector equation), define the NF-dimensional stacked vector $\bar{\mathbf{b}}$ as (recall: F is the number of pricing factors)

$$\underset{(NF \times 1)}{\bar{\mathbf{b}}} \equiv \begin{bmatrix} \underset{(F \times 1)}{\bar{\mathbf{b}}_1} \\ \underset{(F \times 1)}{\bar{\mathbf{b}}_2} \\ \vdots \\ \underset{(F \times 1)}{\bar{\mathbf{b}}_N} \end{bmatrix}. \tag{4.1}$$

Then it is straightforward to show that (3.10), together with the initial condition $\bar{\mathbf{b}}_1 = \boldsymbol{\delta} \equiv (-1, 0, \ldots, 0)'$, can be written as[2]

$$\text{(QVE for } \bar{\mathbf{b}}) \qquad \underset{(NF \times NF)}{\mathbf{M}} \underset{(NF \times 1)}{\bar{\mathbf{b}}} = \underset{(NF \times 1)}{\mathbf{d}} - \gamma \underset{(NF \times 1)}{\mathbf{g}(\bar{\mathbf{b}})}, \tag{4.2}$$

where

$$\underset{(NF \times NF)}{\mathbf{M}} \equiv \begin{bmatrix} \mathbf{I}_F & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ -\underset{(F \times F)}{\mathbf{\Phi}'} & \mathbf{I}_F & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & -\underset{(F \times F)}{\mathbf{\Phi}'} & \mathbf{I}_F & \ldots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & -\underset{(F \times F)}{\mathbf{\Phi}'} & \mathbf{I}_F \end{bmatrix}, \tag{4.3}$$

$$\underset{(NF \times 1)}{\mathbf{d}} \equiv \underset{(N \times 1)}{\mathbf{1}} \otimes \underset{(F \times 1)}{\boldsymbol{\delta}} \quad \text{(where } \boldsymbol{\delta} \equiv (-1, 0, \ldots, 0)' \text{ is the initial condition for } \bar{\mathbf{b}}_1), \tag{4.4}$$

---

[2] Incidentally, there is a corresponding QVE for risk premia. Appendix A shows that risk premia are linear in the factor vector (i.e., affine with zero intercepts). If $\mathbf{h}_n$ is the vector of factor loadings for the risk premium on bonds of maturity n, It is related to the bond price factor loadings $\bar{\mathbf{b}}_n$ as (see (A1.3))

$$\underset{(F \times 1)}{\mathbf{h}_n} = \mathbf{\Phi}'\bar{\mathbf{b}}_{n-1} - \bar{\mathbf{b}}_n + \bar{\mathbf{b}}_1 \quad (n = 2, 3, ., N).$$

Form the $NF \times 1$ stacked vector $\mathbf{h}$ of $(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_N)$ and supplement the above equations by $\mathbf{h}_1 = \boldsymbol{\delta} - \bar{\mathbf{b}}_1$. Then $\mathbf{h}$ is related to $\bar{\mathbf{b}}$ as in

$$\mathbf{h} = -\mathbf{M}\bar{\mathbf{b}} + \mathbf{d}.$$

Solving this for $\bar{\mathbf{b}}$ and substituting into (4.2), we obtain the QVE for the risk-premium factor loadings:

$$\text{(QVE for } \mathbf{h}) \qquad \underset{(NF \times 1)}{\mathbf{h}} = \gamma \mathbf{g}\left(-\underset{(NF \times NF)}{\mathbf{M}^{-1}} \underset{(NF \times 1)}{\underbrace{(\mathbf{h} - \mathbf{d})}}\right).$$

The first F equations state that $\mathbf{h}_1 = \mathbf{0}$, which is the initial condition for $\{\mathbf{h}_n\}$ corresponding to the initial condition for $\{\bar{\mathbf{b}}_n\}$ that $\bar{\mathbf{b}}_1 = \boldsymbol{\delta}$.

$$
\underset{(NF \times 1)}{\mathbf{g}(\bar{\mathbf{b}})} \equiv
\begin{bmatrix}
\underbrace{
\begin{array}{c}
\underset{(F \times 1)}{\mathbf{0}} \\[4pt]
\underset{(F \times (N-1))}{\mathbf{S}'} \; \underset{((N-1) \times F)}{\bar{\mathbf{B}}'} \; \underset{(F \times F)}{\mathbf{\Omega}} \; \underset{(F \times 1)}{\bar{\mathbf{b}}_1}
\end{array}
}_{(F \times 1)} \\[20pt]
\vdots \\[10pt]
\underbrace{\underset{(F \times (N-1))}{\mathbf{S}'} \; \underset{((N-1) \times F)}{\bar{\mathbf{B}}'} \; \underset{(F \times F)}{\mathbf{\Omega}} \; \underset{(F \times 1)}{\bar{\mathbf{b}}_{N-1}}}_{(F \times 1)}
\end{bmatrix}
= \underset{(NF \times (N-1)^2)}{\mathbf{P}} \; \underbrace{\mathrm{vec}\left(\underset{((N-1) \times (N-1))}{\bar{\mathbf{B}}' \, \mathbf{\Omega} \, \bar{\mathbf{B}}} \right)}_{((N-1)^2 \times 1)}, \quad
\mathbf{P} \equiv
\underbrace{\begin{bmatrix}
\underset{(F \times (N-1)^2)}{\mathbf{0}} \\[4pt]
\underbrace{\mathbf{I}_{N-1} \otimes \mathbf{S}'}_{((N-1)F \times (N-1)^2)}
\end{bmatrix}}_{(NF \times (N-1)^2)},
\tag{4.5}
$$

$$
\underset{(F \times (N-1))}{\bar{\mathbf{B}}} \equiv
\begin{bmatrix}
\underset{(F \times 1)}{\bar{\mathbf{b}}_1} & \underset{(F \times 1)}{\bar{\mathbf{b}}_2} & \cdots & \underset{(F \times 1)}{\bar{\mathbf{b}}_{N-1}}
\end{bmatrix}.
\tag{4.6}
$$

4.2. The ODE

If arbitrageurs are risk neutral, i.e., if the risk-aversion parameter γ in (3.3) is zero, then the Eq. (4.2) has a unique solution

$$
\bar{\mathbf{b}}^* \equiv \mathbf{M}^{-1} \mathbf{d}.
\tag{4.7}
$$

Although the square matrix $\mathbf{M}$, which is of size NF, can be large, computing $\bar{\mathbf{b}}^*$ is fast because there is a well-known analytical expression for the inverse of the large matrix $\mathbf{M}$.[3] If $\gamma > 0$, because of the quadratic term $\mathbf{g}(\bar{\mathbf{b}})$, there can be multiple solutions if a solution exists. We seek the solution that converges to $\bar{\mathbf{b}}^*$ as $\gamma \downarrow 0$.[4]

To study such a solution for a given value $\gamma > 0$ of the risk-aversion parameter, consider the following system of equations in $\bar{\mathbf{b}}$ indexed by $\tilde{\gamma}$.

$$
\mathbf{f}(\bar{\mathbf{b}}, \tilde{\gamma}) = \mathbf{0}, \quad \mathbf{f}(\bar{\mathbf{b}}, \tilde{\gamma}) \equiv \mathbf{M}\bar{\mathbf{b}} - \mathbf{d} + \tilde{\gamma} \, \mathbf{g}(\bar{\mathbf{b}}).
\tag{4.9}
$$

This system reduces to (4.2) if $\tilde{\gamma} = \gamma$. The function $\mathbf{f} : \mathbb{R}^{NF} \times \mathbb{R} \to \mathbb{R}^{NF}$ is $\mathbf{C}^\infty$. Its Jacobian with respect to $\bar{\mathbf{b}}$ is given by

$$
\frac{\partial \mathbf{f}(\bar{\mathbf{b}}, \tilde{\gamma})}{\partial \bar{\mathbf{b}}'} = \mathbf{M} + \tilde{\gamma} \, \frac{\partial \mathbf{g}(\bar{\mathbf{b}})}{\partial \bar{\mathbf{b}}'}
\tag{4.10}
$$

By construction, $\bar{\mathbf{b}} = \bar{\mathbf{b}}^*$ is the only solution to (4.9) when $\tilde{\gamma} = 0$. At $(\bar{\mathbf{b}}, \tilde{\gamma}) = (\bar{\mathbf{b}}^*, 0)$, the Jacobian equals $\mathbf{M}$ which is invertible. Therefore, by the Classical Implicit Function Theorem, there exist a neighborhood V of $\bar{\mathbf{b}}^*$, an open interval U containing $\tilde{\gamma} = 0$, and a $\mathbf{C}^\infty$ function $\bar{\mathbf{b}}(\tilde{\gamma}) : U \to V$ such that $\bar{\mathbf{b}}(0) = \bar{\mathbf{b}}^*$ and $\mathbf{f}(\bar{\mathbf{b}}(\tilde{\gamma}), \tilde{\gamma}) = \mathbf{0}$ for all $\tilde{\gamma} \in U$. Its derivative $\bar{\mathbf{b}}'(.)$ is given by

$$
\bar{\mathbf{b}}'(\tilde{\gamma}) = -\left[\frac{\partial \mathbf{f}(\bar{\mathbf{b}}(\tilde{\gamma}), \tilde{\gamma})}{\partial \bar{\mathbf{b}}'} \right]^{-1} \frac{\partial \mathbf{f}(\bar{\mathbf{b}}(\tilde{\gamma}), \tilde{\gamma})}{\partial \gamma} = -\underbrace{\left[\mathbf{M} + \tilde{\gamma} \, \frac{\partial \mathbf{g}(\bar{\mathbf{b}}(\tilde{\gamma}))}{\partial \bar{\mathbf{b}}'} \right]^{-1}}_{(NF \times NF)} \underbrace{\mathbf{g}(\bar{\mathbf{b}}(\tilde{\gamma}))}_{(NF \times 1)}.
\tag{4.11}
$$

This is an ODE (ordinary differential equation) in $\tilde{\gamma}$ defined on U. If a given value $\gamma > 0$ of the risk-aversion parameter is in the open interval U, the ODE can be solved on the interval $[0, \gamma]$ with the initial condition $\bar{\mathbf{b}}(0) = \bar{\mathbf{b}}^*$. The value of the solution path at $\tilde{\gamma} = \gamma$, $\bar{\mathbf{b}}(\gamma)$, is the desired solution to the QVE.[5]

---

[3] Here is the expression for $\mathbf{M}^{-1}$:

$$
\mathbf{M}^{-1} =
\begin{bmatrix}
\mathbf{I}_F & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{\Phi}' & \mathbf{I}_F & \mathbf{0} & \cdots & \mathbf{0} \\
(\mathbf{\Phi}')^2 & \mathbf{\Phi}' & \mathbf{I}_F & \cdots & \mathbf{0} \\
\vdots & & \ddots & \ddots & \mathbf{0} \\
(\mathbf{\Phi}')^{N-1} & \cdots & (\mathbf{\Phi}')^2 & \mathbf{\Phi}' & \mathbf{I}_F
\end{bmatrix}.
\tag{4.8}
$$

[4] As discussed in McCafferty and Driskill (1980), multiple equilibria is generic to nonlinear rational expectations models. To cite Greenwood and Vayanos (2014): "If yields are highly sensitive to shocks to the supply risk factor... , then bonds become highly risky for arbitrageurs. Hence, arbitrageurs absorb supply shocks only if they are compensated by large changes in yields, making the high sensitivity of yields to shocks self-fulfilling." They rule out such equilibria and note that the equilibrium they select "is well behaved in the sense that when a [their risk-aversion parameter] converges to zero, it converges to the unique equilibrium that exists for $a = 0$."

[5] As an illustration, consider the case in which $NF = 1$ and $\mathbf{g}(\bar{\mathbf{b}}) = \bar{b}^2$. The QVE (4.2) or (4.9) is a quadratic equation $\tilde{\gamma} \bar{b}^2 + m \bar{b} - d = 0$ and the ODE (4.11) is $\bar{b}' = -\bar{b}^2/(m + 2\tilde{\gamma}\bar{b})$. If $\tilde{\gamma} = 0$, the unique solution is $\bar{b} = \bar{b}^* = d/m$. If $\tilde{\gamma} \neq 0$, then there are two solutions, $(-m \pm \sqrt{m^2 + 4\tilde{\gamma}d})/(2\tilde{\gamma})$, provided $m^2 + 4\tilde{\gamma}d > 0$. The open interval U for $\tilde{\gamma}$ can be taken to be $(-m^2/d, \infty)$ if $d > 0$ and $(-\infty, -m^2/d)$ if $d < 0$. The solution to the quadratic equation that converges to $\bar{b}^*$ as $\tilde{\gamma} \downarrow 0$ is $(-m + \sqrt{m^2 + 4\tilde{\gamma}d})/(2\tilde{\gamma})$, which also is the solution $\bar{b}(\tilde{\gamma})$ to the ODE.

The Jacobian matrix $\frac{\partial \mathbf{g}(\bar{\mathbf{b}}(\tilde{\gamma}))}{\partial \bar{\mathbf{b}}'}$ in (4.11) has a closed-form expression thanks to $\mathbf{g}(\bar{\mathbf{b}}) \equiv \mathbf{P}\,\mathrm{vec}(\bar{\mathbf{B}}'\mathbf{\Omega}\bar{\mathbf{B}})$ being quadratic. It can be written as (recall: $\underset{(F \times (N-1))}{\bar{\mathbf{B}}} = [\bar{\mathbf{b}}_1\ \bar{\mathbf{b}}_2\ \cdots\ \bar{\mathbf{b}}_{N-1}]$ and $\bar{\mathbf{b}}' = [\bar{\mathbf{b}}_1', \ldots, \bar{\mathbf{b}}_{N-1}', \bar{\mathbf{b}}_N']$)

$$\underbrace{\frac{\partial \mathbf{g}(\bar{\mathbf{b}})}{\partial \bar{\mathbf{b}}'}}_{(NF \times NF)} = \underset{(NF \times (N-1)^2)}{\mathbf{P}} \underbrace{\left[\underset{((N-1)^2 \times (N-1)F)}{\mathbf{A}} \quad \underset{((N-1)^2 \times F)}{\mathbf{0}} \right]}_{((N-1)^2 \times NF)} \tag{4.12}$$

where **A** here is

$$\underset{((N-1)^2 \times (N-1)F)}{\mathbf{A}} \equiv \begin{bmatrix} \underbrace{\mathbf{I}_{N-1} \otimes \underset{(1 \times F)}{\bar{\mathbf{b}}_1'}\,\underset{(F \times F)}{\mathbf{\Omega}}}_{((N-1) \times (N-1)F)} \\ \vdots \\ \underbrace{\mathbf{I}_{N-1} \otimes \underset{(1 \times F)}{\bar{\mathbf{b}}_{N-1}'}\,\underset{(F \times F)}{\mathbf{\Omega}}}_{((N-1) \times (N-1)F)} \end{bmatrix} + \mathbf{I}_{N-1} \otimes \underbrace{\underset{((N-1) \times F)}{\bar{\mathbf{B}}'}\,\underset{(F \times F)}{\mathbf{\Omega}}}_{((N-1) \times F)}. \tag{4.13}$$

4.3. An iterative fixed-point algorithm

Albeit not requiring a numerical calculation of the Jacobian matrix $\frac{\partial \mathbf{g}(\bar{\mathbf{b}}(\tilde{\gamma}))}{\partial \bar{\mathbf{b}}'}$, this method can be computationally demanding if the dimension of the matrix to be inverted, $\mathbf{M} + \tilde{\gamma}\,\frac{\partial \mathbf{g}(\bar{\mathbf{b}}(\tilde{\gamma}))}{\partial \bar{\mathbf{b}}'}$, is large. An obvious alternative is to utilize the fixed-point iteration

$$\bar{\mathbf{b}}^{(\tau+1)} = \mathbf{M}^{-1}[\mathbf{d} - \gamma \mathbf{g}(\bar{\mathbf{b}}^{(\tau)})], \quad \tau = 0, 1, 2, \ldots \tag{4.14}$$

starting from $\bar{\mathbf{b}}^{(0)} = \bar{\mathbf{b}}^*$. This alternative method, with no need to do numerical matrix inversion, is faster by orders of magnitude. A sufficient condition under which this iteration converges to the desired solution $\bar{\mathbf{b}}(\gamma)$ to the QVE is that both $\mathbf{\Phi}$ and $\mathbf{\Omega}$ be nonnegative (see Poloni (2013, Theorem 4)).

5. An illustrative example

The model parameters are: γ (the arbitrageurs' risk-aversion parameter in (3.3)), N (the maximum maturity), (**c, Φ, Ω**) (the VAR parameters describing the dynamics for the pricing factor vector $\mathbf{f}_t$, see (3.5)), and **C** (the matrix connecting supply factors $\boldsymbol{\beta}_t$ to the maturity structure $(s_t^{(2)}, \ldots, s_t^{(N)})$, see (3.7)). The factor vector $\mathbf{f}_t$ consists of the short rate and the supply factors $\boldsymbol{\beta}_t$, as in (3.6). The previous section was about how to compute the affine coefficients (the coefficients of $\mathbf{f}_t$, $\{\bar{a}_n, \bar{\mathbf{b}}_n\}_{n=1}^N$, in (3.9)) given the model parameters.

With computation tools for an arbitrary number of supply factors in hand, we can allow the matrix **C** in (3.7) to be of full rank. Without loss of generality, we can set $\mathbf{C} = \mathbf{I}_{N-1}$. So the supply factors $\boldsymbol{\beta}_t$ are the maturity structure itself, with the factor vector $\mathbf{f}_t$ in (3.6) given by

$$\underset{(N \times 1)}{\mathbf{f}_t} = \left(y_t^{(1)}, s_t^{(2)}, s_t^{(3)}, \ldots, s_t^{(N)} \right)'. \tag{5.1}$$

Consequently, the supply shocks $(\varepsilon_t^{(2)}, \varepsilon_t^{(3)}, \ldots, \varepsilon_t^{(N)})$ are now shocks to the maturity structure, K (the number of supply factors) equals $N - 1$, and F (the dimension of the factor vector $\mathbf{f}_t$) equals $N\ (= 1 + K)$. The affine function for yields in (3.9) is now

$$y_t^{(n)} = a_n + \mathbf{b}_n' \begin{bmatrix} y_t^{(1)} \\ s_t^{(2)} \\ s_t^{(3)} \\ \vdots \\ s_t^{(N)} \end{bmatrix}, \quad n = 1, 2, \ldots, N. \tag{5.2}$$

If b_{jn} is the jth element of $\mathbf{b}_n$, the profile $(b_{j1}, b_{j2}, \ldots, b_{jN})$ describes the impact response of the entire cross section of the yield curve to a one time impulse in $s_t^{(j)}$, the bond supply for the originating maturity j, with all other bond supplies dated t or earlier held constant. That is,

$$\left(\frac{\partial y_t^{(1)}}{\partial s_t^{(j)}}, \frac{\partial y_t^{(2)}}{\partial s_t^{(j)}}, \ldots, \frac{\partial y_t^{(N)}}{\partial s_t^{(j)}} \right) = (b_{j1}, b_{j2}, \ldots, b_{jN}). \tag{5.3}$$

The example considered in this section has a factor dynamics with the following features: (i) the bond supply shock volatility is the same across maturities (so maturity-specific effects on yields, if any, are not attributable to a supply shock

volatility unique to the maturity), (ii) the maturity structure dynamics arises solely from what we call the "legacy" nature of bonds that an n-period bond becomes an $(n-1)$-period bond in the next period, and (iii) the short rate is exogenous to the maturity structure (so the supply effect on the yield curve is not contaminated by the expectations effect that operates through changes in the future path of the short rate).

5.1. Specifying the factor dynamics

The VAR factor dynamics, $\mathbf{f}_{t+1} = \mathbf{c} + \mathbf{\Phi}\,\mathbf{f}_t + \boldsymbol{\varepsilon}_{t+1}$, $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ (reproduced from (3.5)), is specified as

$$\text{(short rate)} \quad y_t^{(1)} = c_1 + \rho y_{t-1}^{(1)} + \varepsilon_{1t}, \quad \varepsilon_{1t} \sim \mathcal{N}(0, \sigma_1^2), \tag{5.4}$$

$$\text{(maturity structure)} \quad s_t^{(n)} = \begin{cases} c_n + \theta s_{t-1}^{(n+1)} + \varepsilon_{nt}, & \varepsilon_{nt} \sim \mathcal{N}(0, \sigma^2) & \text{if } n = 2, 3, \ldots, N-1, \\ c_N + \varepsilon_{Nt}, & \varepsilon_{Nt} \sim \mathcal{N}(0, \sigma^2) & \text{if } n = N. \end{cases} \tag{5.5}$$

The volatility of supply shocks ($\varepsilon_{2t}, \ldots, \varepsilon_{Nt}$) are uniform across maturities because $\mathrm{Var}(\varepsilon_{nt})$ ($=\sigma^2$) does not depend on $n = 2, 3, \ldots, N$. The short rate $y_t^{(1)}$ is exogenous to the maturity structure, because the short rate shock ε_{1t} is assumed to be uncorrelated with the bond supply shocks and also because the lagged supply of bonds do not enter the short rate equation. The extent of the bond "legacy" property is represented by θ, the fraction of the n-period bond supply that will be left as $(n-1)$-period bonds in the next period. It equals 0 if there is a fully offsetting market operation in the next period, and 1 if there is no such offsetting.[6]

In all calibrations reported below, neither $\mathbf{\Phi}$ nor $\boldsymbol{\Omega}$ have negative entries, so by Poloni (2013, Theorem 4) the fixed-point iteration (4.14) converges to the desired solution to the QVE.

5.2. Calibration

To avoid heavy demand on CPU time, we choose the unit interval to be a quarter, not a month, and set N (the maximum maturity) to 80 quarters (20 years). So the short rate $y_t^{(1)}$ is the 3-month interest rate.

The factor dynamics consists of the AR(1) short-rate Eq. (5.4) and the maturity structure dynamics (5.5). The former is estimated by OLS (ordinary least squares) on the period under Alan Greenspan's chairmanship (1987:Q3 - 2005:Q4).[7] The estimated AR(1) equation is (with standard errors in parentheses, yields here stated in annual percentage points)

$$y_t^{(1)} = \underset{(0.149)}{0.145} + \underset{(0.029)}{0.9632}\, y_{t-1}^{(1)}, \quad \text{S.E.R.} = 0.520, \quad t = 1987\text{:Q3-2005:Q4}. \tag{5.6}$$

This determines ρ and σ_1 in (5.4): $\rho = 0.9632$, and $\sigma_1 = 0.520\%$ (or 0.13% if the interest rates are per quarter). The value of c_1 is set so that the steady-state value, $c_1/(1-\rho)$, equals the sample mean over 1987:Q3-2005:Q4 of $y_t^{(1)}$.

Regarding the maturity structure dynamics (5.5), we pick the intercept terms ($c_2, \ldots, c_N$) so that the steady-state value of $s_t^{(n)}$ is $1/N$ for $n = 2, \ldots, N$.[8] For now we assume supply shocks are local in that there is no correlation between supply shocks of different maturities. Therefore the supply shock variance matrix is peripheral: $\mathrm{Var}(\varepsilon_{2t}, \ldots, \varepsilon_{Nt}) = \sigma^2 \mathbf{I}_{N-1}$. We set the supply shock volatility σ (i.e., the standard deviation of ε_{nt}, common to all maturities $n = 2, 3, \ldots, N$) to 0.005 ($= 0.5\%$).

---

[6] The implied restrictions on the VAR parameters ($\mathbf{\Phi}$, $\boldsymbol{\Omega}$) are as follows.

$$\underset{(N \times N)}{\mathbf{\Phi}} = \begin{bmatrix} \rho & \underset{(1 \times (N-1))}{\mathbf{0}'} \\ \underset{((N-1) \times 1)}{\mathbf{0}} & \underset{((N-1) \times (N-1))}{\tilde{\mathbf{\Phi}}} \end{bmatrix}, \quad \underset{(N \times N)}{\boldsymbol{\Omega}} = \begin{bmatrix} \sigma_1^2 & \underset{(1 \times (N-1))}{\mathbf{0}'} \\ \underset{((N-1) \times 1)}{\mathbf{0}} & \underset{((N-1) \times (N-1))}{\tilde{\boldsymbol{\Omega}}} \end{bmatrix},$$

where $\tilde{\mathbf{\Phi}}$ is given by

$$\underset{((N-1) \times (N-1))}{\tilde{\mathbf{\Phi}}} = \begin{bmatrix} 0 & \theta & 0 & 0 & \cdots & 0 \\ 0 & 0 & \theta & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \theta & 0 \\ 0 & 0 & \cdots & \cdots & 0 & \theta \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}.$$

[7] The yield curve, and the short rate in particular, are computed from Gurkaynak et al. (2007) as follows. The parametrized function used to fit the yields of various observed maturities is given as Eq. (22) in Gurkaynak et al. (2007). The daily parameter values can be downloaded from the Federal Reserve Board website. From this we can calculate $y_t^{(n)}$ for any maturity n. We take the yield for the quarter to be for the last business day of the quarter. For the short rate we use the value implied by the parametrized yield curve at $n = 1$ (3 months). This short rate series is very similar to the constant-maturity 3-month rate available from Federal Reserve Board's Table H-15.

[8] Let $\underset{((N-1) \times 1)}{\tilde{\mathbf{s}}_t} \equiv (s_t^{(2)}, \ldots, s_t^{(N)})'$ be the maturity structure, $\tilde{\mathbf{c}} \equiv (c_2, \ldots, c_{N-1}, c_N)'$, and let $\tilde{\mathbf{\Phi}}$ be as in footnote 6. Then the steady-state value of $\tilde{\mathbf{s}}_t$, call it $\tilde{\mathbf{s}}$, satisfies: $(\mathbf{I} - \tilde{\mathbf{\Phi}})\tilde{\mathbf{s}} = \tilde{\mathbf{c}}$.
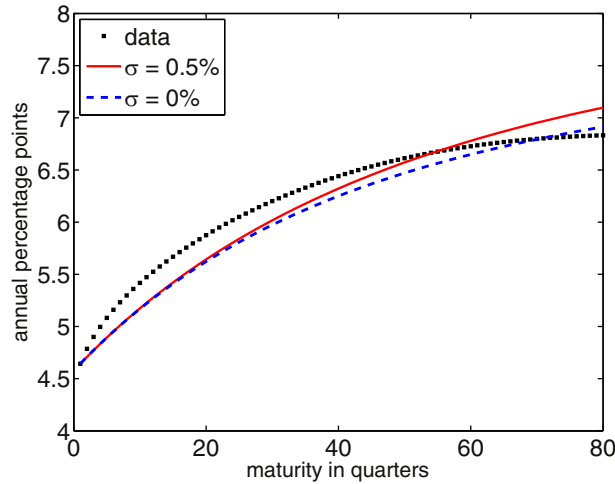
Fig. 1. Average yield curve. *Note:* $(\gamma, \theta) = (42, 1)$ where γ is the risk-aversion parameter and θ is the extent of legacy. The graph with $\theta = 0$ is similar. The supply shock volatility σ is 0.005 (0.5%) for the solid line and 0% for the dashed line.

In the response analysis below, the size of the one-time impulse in bond supply $s_t^{(j)}$ is 1 percentage point, or twice the supply shock volatility σ, which formalizes the notion of "large scale" bond supply changes.

This leaves two parameters, γ (the risk-aversion parameter) and θ (the legacy parameter that controls the maturity structure dynamics). The parameter θ, which controls the short-run maturity structure dynamics, has only modest effects on the steady-state yield curve.[9] We can therefore calibrate γ to the steady-state yield curve almost regardless of the value of θ. We set $\gamma = 42$ because, as shown in Fig. 1, the model-implied steady-state yield curve with $(\gamma, \theta) = (42, 1)$ and the corresponding sample mean for the sample period (1987:Q3 - 2005:Q4) are close to each other.[10] (Ignore the dashed line in the figure for now.)

5.3. Response of risk premia to local supply shocks

We are now ready to describe the effect of bond supply shocks. We first examine the effect on risk premia before turning to yields, in order to exploit the formula (3.13) which connects risk premia to yields. This is also the route taken by Greenwood and Vayanos (2014) for their model of a single supply factor. It is straightforward to show (see (A1.3)) that risk premia are linear (affine with zero intercepts) in $\mathbf{f}_t$. Let h_{jn} be the coefficient of the jth element of the coefficient vector $\mathbf{h}_n$.[11] The (cross-section) profile of impact responses $(\frac{\partial \mathrm{rp}_t^{(1)}}{\partial s_t^{(j)}}, \frac{\partial \mathrm{rp}_t^{(2)}}{\partial s_t^{(j)}}, \ldots, \frac{\partial \mathrm{rp}_t^{(N)}}{\partial s_t^{(j)}})$ of the term structure of risk premia to a one-time impulse in bond supply originating in maturity j ($j = 2, 3, \ldots, N$) is given by the profile $(h_{j1}, h_{j2}, \ldots, h_{jN})$.

Fig. 2 graphs the risk-premium profile for three originating maturities, 20 quarters (5 years), 40 quarters (10 years), and 80 quarters (20 years). The impulse is a 1 percentage point increase in bond supply $s_t^{(j)}$. The profile shown is only for $\theta = 1$, because (in sharp contrast to the yield profiles to be discussed below) the risk-premium profiles for other values of θ, $0 \leq \theta < 1$, are lower but only slightly so under the calibrated parameter values. There are two notable features:

- The profile shifts up as the originating maturity gets longer. That is, $\frac{\partial \mathrm{rp}_t^{(n)}}{\partial s_t^{(j)}}$ is increasing in j for each $n = 2, 3, \ldots, N$.
  An increased supply of bonds of maturity j ($= 2, 3, \ldots, N$), which is to be purchased by arbitrageurs, makes the bond portfolio held by arbitrageurs more sensitive to the short rate shock. Thus the portfolio's volatility (the $\mathrm{Var}(R_{t+1})$ in the first-order conditions (3.4)) increases, which raises the risk premium on bonds of all maturities. This effect is stronger if the originating maturity (j) of the additional bonds supplied is longer, because the portfolio's volatility increase is greater.

- The profiles are upward-sloping. That is, $\frac{\partial \mathrm{rp}_t^{(n)}}{\partial s_t^{(j)}}$ is increasing in n as well. The risk premium needed to compensate arbitrageurs for the portfolio's volatility increase is greater for longer-term bonds.[12]

---

[9] Let $\boldsymbol{\mu}$ be the unconditional mean of the VAR (3.5), so $(\mathbf{I} - \boldsymbol{\Phi})\boldsymbol{\mu} = \mathbf{c}$. The steady-state yield curve is given by $y_t^{(n)} = a_n + \mathbf{b}_n'\boldsymbol{\mu}$, $n = 1, 2, \ldots, N$.

[10] The value of γ calibrated by Greenwood and Vayanos (2014), where bond returns are stated at annual rates as here, ranges from 42 to 64.

[11] Footnote 2 shows how $\{\mathbf{h}_n\}$ are related to the yield factor loadings $\{\mathbf{b}_n\}$.

[12] This feature carries over to the case of a single supply factor that increases the average duration, see Greenwood and Vayanos (2014, Proposition 2).
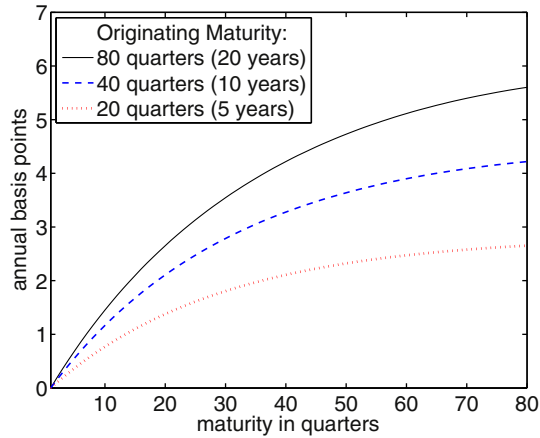
Fig. 2. Response of risk premia to supply shocks. *Note:* The impact response of the term structure of risk premia. The impulse is a 1 percentage point increase in the share of bonds of the originating maturity. $(\gamma, \theta) = (42, 1)$ where γ is the risk-aversion parameter and θ is the extent of legacy. The graph with $\theta = 0$ is similar. The supply shock volatility σ is 0.005 (0.5%).

5.4. Response of the yield curve to local supply shocks

These two features about the cross-section profile of the risk-premium responses can be utilized to characterize the yield-curve responses. We proceed in two steps. First, differentiating both sides of the formula (3.13) with respect to $s_t^{(j)}$ and noting that (i) the shock does not affect current and future short rates and (ii) subsequent responses, $\frac{\partial \operatorname{rp}_{t+k}^{(n)}}{\partial s_t^{(j)}}$ $(k = 1, 2, \ldots)$, are known at date t (so the conditional expectations operator E_t can be dropped), we obtain

$$\frac{\partial y_t^{(n)}}{\partial s_t^{(j)}} = \frac{1}{n}\left[\frac{\partial \operatorname{rp}_t^{(n)}}{\partial s_t^{(j)}} + \frac{\partial \operatorname{rp}_{t+1}^{(n-1)}}{\partial s_t^{(j)}} + \cdots + \frac{\partial \operatorname{rp}_{t+n-2}^{(2)}}{\partial s_t^{(j)}} \right], \quad n = 2, 3, \ldots, N. \tag{5.7}$$

That is, the impact response of the n-period yield to a supply shock originating in maturity j is the average of the *longitudinal* profile of risk-premium responses over the life of the n-period bond.

Second, the longitudinal risk-premium profile, in turn, is linked to the cross-section risk-premium profiles according to the following formula:

$$\left(\underbrace{\frac{\partial \operatorname{rp}_t^{(n)}}{\partial s_t^{(j)}}, \frac{\partial \operatorname{rp}_{t+1}^{(n-1)}}{\partial s_t^{(j)}}, \ldots, \frac{\partial \operatorname{rp}_{t+n-2}^{(2)}}{\partial s_t^{(j)}}}_{n-1 \text{ elements}} \right) = \left(\frac{\partial \operatorname{rp}_t^{(n)}}{\partial s_t^{(j)}}, \theta \frac{\partial \operatorname{rp}_t^{(n-1)}}{\partial s_t^{(j-1)}}, \theta^2 \frac{\partial \operatorname{rp}_t^{(n-2)}}{\partial s_t^{(j-2)}}, \ldots \right). \tag{5.8}$$

This holds because fraction θ is left in the market one period after an increase in the bond supply (so, for example, $\frac{\partial \operatorname{rp}_{t+1}^{(n-1)}}{\partial s_t^{(j)}}$ equals $\theta \frac{\partial \operatorname{rp}_{t+1}^{(n-1)}}{\partial s_{t+1}^{(j-1)}}$) and also because the impact and subsequent responses are timeless (so, for example, $\frac{\partial \operatorname{rp}_{t+1}^{(n-1)}}{\partial s_{t+1}^{(j-1)}}$ equals $\frac{\partial \operatorname{rp}_t^{(n-1)}}{\partial s_t^{(j-1)}}$). Due to this link, the longitudinal risk-premium profile has the following two features.

(a) The profile shifts up with originating maturity j, because (as noted above) $\frac{\partial \operatorname{rp}_t^{(n)}}{\partial s_t^{(j)}}$ is increasing in j.

(b) The profile is a declining sequence, because $\theta \leq 1$ and because (as noted above) $\frac{\partial \operatorname{rp}_t^{(n)}}{\partial s_t^{(j)}}$ declines with n as well as with j.[13]

The yield impact response profile $\left(\frac{\partial y_t^{(1)}}{\partial s_t^{(j)}}, \frac{\partial y_t^{(2)}}{\partial s_t^{(j)}}, \ldots, \frac{\partial y_t^{(N)}}{\partial s_t^{(j)}} \right)$ can be calculated from cross-section risk-premium profiles for various j's, some of which are shown in Fig. 2, via (5.7) and (5.8). Fig. 3 displays the profile for the three originating maturities, for $\theta = 1$. Thanks to the two features, (a) and (b) above, about risk-premium profiles, the yield profiles have the following features that hold for any given θ between 0 and 1.

---

[13] For example, the sequence's first element is greater than its second element because: $\frac{\partial \operatorname{rp}_t^{(n)}}{\partial s_t^{(j)}} > \frac{\partial \operatorname{rp}_t^{(n-1)}}{\partial s_t^{(j)}} > \frac{\partial \operatorname{rp}_t^{(n-1)}}{\partial s_t^{(j-1)}} \geq \theta \frac{\partial \operatorname{rp}_t^{(n-1)}}{\partial s_t^{(j-1)}}$.
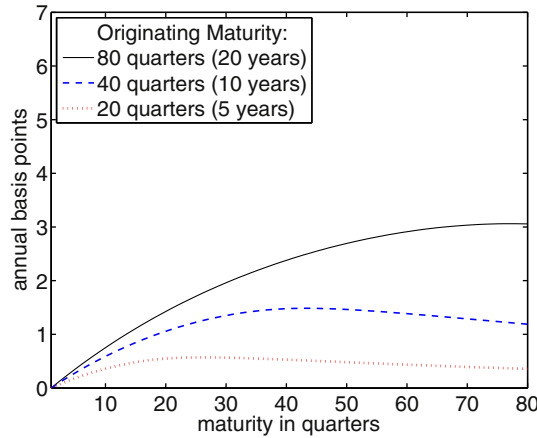
Fig. 3. Response of yields to supply shocks. *Note:* The impact response of the yield curve. The impulse is a 1 percentage point increase in the share of bonds of the originating maturity. $(\gamma, \theta) = (42, 1)$ where γ is the risk-aversion parameter and θ is the extent of legacy. The graph with $\theta = 0$ is very close to the horizontal line. The supply shock volatility σ is 0.005 (0.5%).

- For each originating maturity, the profile is greater for risk premia than for yields. That is, $\frac{\partial y_t^{(n)}}{\partial s_t^{(j)}} < \frac{\partial \, \mathrm{rp}_t^{(n)}}{\partial s_t^{(j)}}$ for all n, j $= 2, 3, \ldots, N$. Even with $\theta = 1$, the profiles shown in Fig. 2 lie above the corresponding profiles shown in Fig. 3 for each originating maturity. The reason is that the longitudinal profile $\left(\frac{\partial \, \mathrm{rp}_t^{(n)}}{\partial s_t^{(j)}}, \frac{\partial \, \mathrm{rp}_{t+1}^{(n-1)}}{\partial s_t^{(j)}}, \ldots, \frac{\partial \, \mathrm{rp}_{t+n-2}^{(2)}}{\partial s_t^{(j)}} \right)$, whose average equals $\frac{\partial y_t^{(n)}}{\partial s_t^{(j)}}$ by (5.7), is a declining sequence, as noted in (b) above.[14]

- As in Fig. 2 for risk premia, the yield profile shifts up as the originating maturity gets longer. That is, $\frac{\partial y_t^{(n)}}{\partial s_t^{(j)}}$ is increasing in j for $n = 2, 3, \ldots, N$. This follows directly from (a).

- Unlike the risk-premium profiles shown in Fig. 2, the yield profiles are hump-shaped around the originating maturity. The reason is that the effect of a supply shock at originating maturity j lasts only up to $j - 1$ periods. For example, an increase in $s_t^{(2)}$ ceases to exert yield-curve effects in the next period $t + 1$ when two-period bonds turn into one-period securities. More generally, for bonds whose maturity n is greater than j, the longitudinal profile in (5.8) is truncated, as in

$$\left(\underbrace{\frac{\partial \, \mathrm{rp}_t^{(n)}}{\partial s_t^{(j)}}, \ldots, \frac{\partial \, \mathrm{rp}_{t+n-2}^{(2)}}{\partial s_t^{(j)}}}_{n-1 \text{ elements}} \right) = \left(\underbrace{\frac{\partial \, \mathrm{rp}_t^{(n)}}{\partial s_t^{(j)}}, \ldots, \theta^{j-2} \frac{\partial \, \mathrm{rp}_t^{(n-j+2)}}{\partial s_t^{(2)}}}_{j-1 \text{ elements}}, \underbrace{0, \ldots, 0}_{n-j \text{ elements}} \right). \tag{5.9}$$

There are two opposing effects on the average of this profile as n gets longer given j. One is a dilution effect due to the truncation after $j - 1$ periods. The other is that the profile before truncation is higher because the risk-premium response increases with n. The former effect kicks in as n gets past j and, under the calibrated parameter values, starts dominating the latter effect.

The preceding discussion is for any given value of θ, $0 \leq \theta \leq 1$. We noted earlier for risk premia that response profiles are not very sensitive to θ under the calibrated parameter values. This is not the case for yields because, as (5.8) makes it clear, the current yield response is a weighted sum of current and future risk-premium responses discounted by θ. Indeed, our calculation results (not reported here) show that the yield response profiles, shown in Fig. 3 for $\theta = 1$, fall rapidly to the horizontal axis as θ is decreased from 1 to 0.

We repeated the preceding calculations for smaller values for γ (the risk-aversion parameter). The risk premium and yield responses get smaller, roughly in proportion to the decline in γ. The yield responses continue to be hump-shaped. More detailed results are not reported here because the model had difficulty generating a curvature in the yield curve that is even remotely close to the curvature shown in Fig. 1.

5.5. Allowing for correlation in supply shocks across maturities

So far, our calculation points to modest supply effects. As Fig. 3 indicates, a one percentage point impulse raises the yield by only about 3 basis points at the long end of the maturity spectrum, where the supply effect should be greatest. The

---

[14] This feature carries over to the case of a single supply factor that increases the average duration, see Greenwood and Vayanos (2014, Proposition 3).
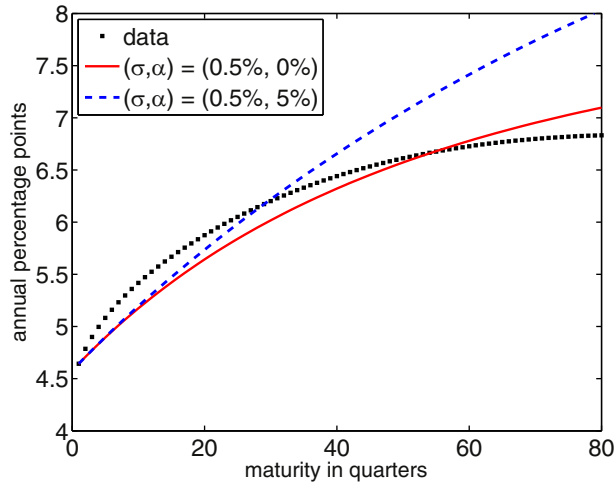
Fig. 4. Average Yield Curve with Correlated Supply Shocks. *Note:* $(\gamma, \theta) = (42, 1)$ where γ is the risk-aversion parameter and θ is the extent of legacy. The graph with $\theta = 0$ is similar. $(\sigma, \alpha) = (0.5\%, 0\%)$ for the solid line and $(\sigma, \alpha) = (0.5\%, 5\%)$ for the dashed line, where σ is the supply shock volatility and α is the correlation coefficient between supply shocks of different maturities.
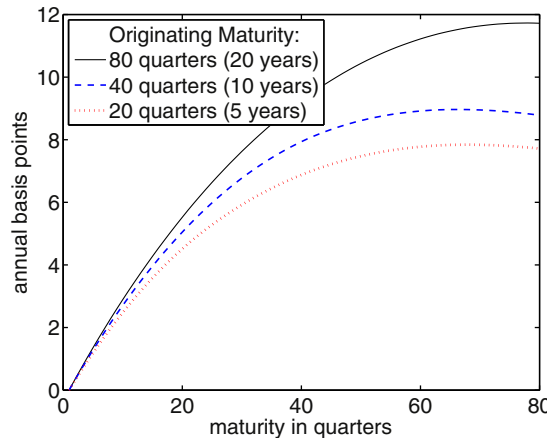

Fig. 5. Response of yields to correlated supply shocks. *Note:* The impact response of the yield curve. The shock is a 1 percentage point increase in the share of bonds of the originating maturity. $(\gamma, \theta) = (42, 1)$ where γ is the risk-aversion parameter and θ is the extent of legacy. The graph with $\theta = 0$ is very close to the horizontal line. $(\sigma, \alpha) = (0.5\%, 5\%)$ where σ is the supply shock volatility and α is the correlation coefficient between supply shocks of different maturities.

steady-state yield curve, too, is only modestly affected, as indicated in Fig. 1 where the dashed line − the yield curve that would prevail if the supply shock volatility σ were zero − lies only slightly below the curve with $\sigma = 0.5\%$.

The supply effect is necessarily modest if supply shocks can be easily diversified away in the bond portfolio held by the arbitrageurs. To examine the extent of diversification, we now consider supply shocks that are correlated across maturities. For the supply shock variance matrix $\text{Var}(\varepsilon_{2t}, \ldots, \varepsilon_{Nt})$, instead of assuming it to be peripheral ($\sigma^2 \mathbf{I}_{N-1}$), we allow it to have non-zero off-diagonal elements. We could entertain any pattern of correlation, but here we consider an equi-correlation matrix where the off-diagonal elements are $\alpha\sigma^2$. Thus the correlation coefficient between any pair of shocks at two different maturities is α.

This affects the yield affine coefficients ($\{a_n, \mathbf{b}_n\}_{n=1}^N$ in (5.2)), and hence the stead-state yield curve $\{a_n + \mathbf{b}_n'\boldsymbol{\mu}\}_{n=1}^N$ (where $\boldsymbol{\mu}$ is the long-run value of the factor vector $\mathbf{f}_t$). Fig. 4 describes how correlated supply shocks affect the steady-state yield curve. The dashed line, which in Fig. 1 was the model-implied steady-state yield curve for the case of no supply shocks ($\sigma = 0$), is now for the case of $(\sigma, \alpha) = (0.5\%, 5\%)$. Thus the difference between the solid line (which is for $(\sigma, \alpha) = (0.5\%, 0\%)$) and the dashed line in Fig. 4 is a measure of the long-run effect on the yield curve of an increase in the correlation coefficient α for the bond supply shocks from 0% to 5%.

The impact yield responses are affected as well, for two reasons. First, as just noted, the factor loadings $\{\mathbf{b}_n\}$ are different in value. Under the calibrated parameter values, an increase in the correlation α raises the factor loadings. Second, a one-unit increase in bond supply $s_t^{(j)}$ brings about an increase of size α in all other maturities, whose effect needs to be accounted

for. More precisely, the impact response of $y_t^{(n)}$ to a bond supply shock at maturity j now equals

$$\frac{\partial y_t^{(n)}}{\partial s_t^{(j)}} + \alpha \sum_{k \neq j} \frac{\partial y_t^{(n)}}{\partial s_t^{(k)}} = b_{jn} + \alpha \sum_{k \neq j} b_{kn}. \tag{5.10}$$

Fig. 5 displays the yield response profiles for $(\sigma, \alpha) = (0.5\%, 5\%)$. Comparing this figure with Fig. 3 where $(\sigma, \alpha) = (0.5\%, 0\%)$, we see that the responses are about three times as large. The hump is less pronounced here because the yield response now includes effects from other maturities (as indicated in (5.10)).

6. Conclusion and agenda for future research

This paper has developed tools for computing a mapping from the maturity structure to equilibrium bond prices for the Vayanos-Vila model. No restrictions are imposed on the VAR representation of the maturity structure dynamics.

It should now be possible to estimate the Vayanos–Vila model. The system would consist of two blocks. One is the VAR equations for pricing factors including the maturity structure. The other is a set of equations relating yields to the pricing factors. The mapping provides cross-equation restrictions: the coefficients in the second block are functions of both the VAR parameters from the first block and the risk aversion parameter. It would thus be possible to identify the risk-aversion parameter from information on yields and the maturity structure.

Appendix A. Derivation of factor-loading equations (3.10) and (3.11)

We proceed in two steps:

(a) Under the VAR dynamics (3.5) and the conjecture of the affine bond pricing (3.9), derive the arbitrageurs' FOCs (first-order conditions) that involve the affine coefficients $(\bar{a}_n, \bar{\mathbf{b}}_n)$.
(b) Impose on the FOCs the bond market equilibrium conditions, $z_t^{(n)} = s_t^{(n)}$ for $n = 2, 3, \ldots, N$, to derive a set of equations for the affine coefficients.

Regarding (a), the FOCs for n-period bonds, reproduced from the text and to be derived below, are

$$\mathrm{E}_t \underbrace{\left[\underbrace{\frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}}}_{\text{holding-period return}} \right] - \underbrace{\frac{1 - P_t^{(1)}}{P_t^{(1)}}}_{\text{short rate}}}_{\text{risk premium on } n\text{-period bonds}} = \gamma \frac{1}{2} \frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}}, \quad n = 2, 3, \ldots, N. \tag{A1.1}$$

We can rewrite the risk premium (the left-hand side of (A1.1)) and $\frac{1}{2} \frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}}$ (on the right-hand side of (A1.1)) as

$$\mathrm{E}_t \left[\frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}} \right] - \frac{1 - P_t^{(1)}}{P_t^{(1)}} \approx \mathrm{E}_t \left(\log P_{t+1}^{(n-1)} \right) - \log P_t^{(n)} + \frac{1}{2} \bar{\mathbf{b}}_{n-1}' \boldsymbol{\Omega} \bar{\mathbf{b}}_{n-1} - y_t^{(1)} \tag{A1.2}$$

$$\text{(see below for derivation)}$$

$$= \bar{a}_{n-1} + \bar{\mathbf{b}}_{n-1}'(\mathbf{c} + \boldsymbol{\Phi} \mathbf{f}_t) - \bar{a}_n - \bar{\mathbf{b}}_n' \mathbf{f}_t + \frac{1}{2} \bar{\mathbf{b}}_{n-1}' \boldsymbol{\Omega} \bar{\mathbf{b}}_{n-1} + \bar{a}_1 + \bar{\mathbf{b}}_1' \mathbf{f}_t \quad \text{(by (3.5) and (3.9))}$$

$$= \left(\bar{a}_{n-1} + \bar{\mathbf{b}}_{n-1}' \mathbf{c} - \bar{a}_n + \frac{1}{2} \bar{\mathbf{b}}_{n-1}' \boldsymbol{\Omega} \bar{\mathbf{b}}_{n-1} + \bar{a}_1 \right) + \left(\bar{\mathbf{b}}_{n-1}' \boldsymbol{\Phi} - \bar{\mathbf{b}}_n' + \bar{\mathbf{b}}_1' \right) \mathbf{f}_t, \tag{A1.3}$$

$$\frac{1}{2} \frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}} \approx \bar{\mathbf{b}}_{n-1}' \boldsymbol{\Omega} \left(\bar{\mathbf{b}}_1 z_t^{(2)} + \cdots + \bar{\mathbf{b}}_{N-1} z_t^{(N)} \right), \quad \text{(see below for derivation)} \tag{A1.4}$$

for $n = 2, 3, \ldots, N$. Eqs. (A1.2) and (A1.4), while exact in continuous time, are only approximations in discrete time.

Turning to (b), upon the imposition of the market equilibrium conditions, the expression (A1.4) for $\frac{1}{2}\frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}}$ can be written as

$$
\frac{1}{2}\frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}} \approx \bar{\mathbf{b}}_{n-1}' \, \boldsymbol{\Omega} \, \left(\bar{\mathbf{b}}_1 s_t^{(2)} + \cdots + \bar{\mathbf{b}}_{N-1} s_t^{(N)} \right) \qquad \text{(by replacing } z_t^{(i)} \text{ by } s_t^{(i)}, i = 2, \ldots, N)
$$

$$
= \bar{\mathbf{b}}_{n-1}' \, \boldsymbol{\Omega} \, \begin{bmatrix} \bar{\mathbf{b}}_1 & \bar{\mathbf{b}}_2 & \cdots & \bar{\mathbf{b}}_{N-1} \end{bmatrix} \begin{bmatrix} s_t^{(2)} \\ s_t^{(3)} \\ \vdots \\ s_t^{(N)} \end{bmatrix}
$$

$$
= \bar{\mathbf{b}}_{n-1}' \, \boldsymbol{\Omega} \, \begin{bmatrix} \bar{\mathbf{b}}_1 & \bar{\mathbf{b}}_2 & \cdots & \bar{\mathbf{b}}_{N-1} \end{bmatrix} \mathbf{S}\, \mathbf{f}_t. \qquad \text{(by (3.7))} \tag{A1.5}
$$

Substituting (A1.3) and (A1.5) into (A1.1), we obtain a set of equations that is affine in $\mathbf{f}_t$. This has to hold for any $\mathbf{f}_t$. Setting both the constant term and the coefficients of $\mathbf{f}_t$ to zero, we obtain the system of Eqs., (3.10) and (3.11) in the text, for affine factor loadings $(\bar{a}_n, \bar{\mathbf{b}}_n)$.

A1. Derivation of the FOCs, (A1.1)

The decision problem of arbitrageurs is described by (3.3) with (3.2) of the text. If ϕ_t is the Lagrange multiplier for the constraint $\sum_{n=1}^{N} z_t^{(n)} = 1$, the FOCs (first-order conditions) are

$$
\mathrm{E}_t\left[\frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}} \right] - \gamma \, \frac{1}{2}\frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}} = \phi_t, \qquad n = 1, 2, \ldots, N, \tag{A1.6}
$$

$$
\sum_{n=1}^{N} z_t^{(n)} = 1. \tag{A1.7}
$$

Since $P_{t+1}^{(0)} = 1$, the holding period return on the 1-period bond is $(1 - P_t^{(1)})/P_t^{(1)}$, which is known in date t. Hence $\frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}} = 0$ for $n = 1$ and the FOC for $n = 1$ is: $(1 - P_t^{(1)})/P_t^{(1)} = \phi_t$. Substituting this into the rest of the FOCs in (A1.6), we obtain (A1.1).

A2. Derivation of (A1.2) and (A1.4)

This can be done easily if we show the following.

- Show: $(1 - P_t^{(1)})/P_t^{(1)} \approx y_t^{(1)}$. This immediately follows from the approximation $1/x \approx 1 - \log(x)$ for $x \approx 1$ and the relation $y_t^{(1)} \equiv -\log P_t^{(1)}$.
- Let $p_t^{(n)} \equiv \log P_t^{(n)}$. Show: the distribution conditional on date t of $p_{t+1}^{(n-1)} - p_t^{(n)}$ is normal with mean $\bar{a}_{n-1} + \bar{\mathbf{b}}_{n-1}'(\mathbf{c} + \boldsymbol{\Phi}\mathbf{f}_t) - \bar{a}_n - \bar{\mathbf{b}}_n'\mathbf{f}_t$ and variance $\bar{\mathbf{b}}_{n-1}'\boldsymbol{\Omega}\bar{\mathbf{b}}_{n-1}$ for $n = 2, 3, \ldots N$. This follows immediately from

$$
p_{t+1}^{(n-1)} - p_t^{(n)} = \bar{a}_{n-1} + \bar{\mathbf{b}}_{n-1}'(\mathbf{c} + \boldsymbol{\Phi}\mathbf{f}_t) - \bar{a}_n - \bar{\mathbf{b}}_n'\mathbf{f}_t + \bar{\mathbf{b}}_{n-1}'\boldsymbol{\varepsilon}_{t+1}, \tag{A1.8}
$$

which can be obtained by combining (3.5) and (3.9).

- Show:

$$
\mathrm{E}_t\left[\frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}} \right] = \mathrm{E}_t\left[\exp(p_{t+1}^{(n-1)} - p_t^{(n)}) \right] - 1 \approx \mathrm{E}_t\left(p_{t+1}^{(n-1)} - p_t^{(n)} \right) + \frac{1}{2}\operatorname{Var}_t\left(p_{t+1}^{(n-1)} - p_t^{(n)} \right) \tag{A1.9}
$$

for $n = 2, 3, \ldots, N$. This can be obtained easily from two well-known formulas. One is that, for a normally distributed random variable X, $\mathrm{E}[\exp(X)] = \exp[\mathrm{E}(X) + \frac{1}{2}\operatorname{Var}(X)]$, and the other is $\exp(x) \approx 1 + x$ for $x \approx 0$. In the former, we can set $X = p_{t+1}^{(n-1)} - p_t^{(n)}$ because, as shown above, $p_{t+1}^{(n-1)} - p_t^{(n)}$ is normally distributed conditional on date t information. In the latter formula, set $x = \mathrm{E}_t\left(p_{t+1}^{(n-1)} - p_t^{(n)} \right) + \frac{1}{2}\operatorname{Var}_t\left(p_{t+1}^{(n-1)} - p_t^{(n)} \right)$.

- Show:

$$
\operatorname{Var}_t(R_{t+1}) \approx \mathbf{d}_t'\boldsymbol{\Omega}\mathbf{d}_t \quad \text{where} \quad \mathbf{d}_t \equiv \sum_{k=2}^{N} \bar{\mathbf{b}}_{k-1}\, z_t^{(k)}. \qquad \text{So} \quad \frac{1}{2}\frac{\partial \operatorname{Var}_t(R_{t+1})}{\partial z_t^{(n)}} = \bar{\mathbf{b}}_{n-1}'\boldsymbol{\Omega}\mathbf{d}_t. \tag{A1.10}
$$

To derive this, we again use the approximation $\exp(x) \approx 1 + x$ to derive

$$
\frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}} = \exp\left(p_{t+1}^{(n-1)} - p_t^{(n)} \right) - 1 \approx p_{t+1}^{(n-1)} - p_t^{(n)}. \tag{A1.11}
$$

Multiply both sides by $z_t^{(n)}$ and sum over $n = 1, 2, \ldots, N$ to obtain

$$R_{t+1} \equiv \sum_{n=1}^{N} \frac{P_{t+1}^{(n-1)} - P_t^{(n)}}{P_t^{(n)}} z_t^{(n)} \approx \sum_{n=1}^{N} \left(p_{t+1}^{(n-1)} - p_t^{(n)} \right) z_t^{(n)}$$

$$= -p_t^{(1)} z_t^{(1)} + \sum_{n=2}^{N} \left(p_{t+1}^{(n-1)} - p_t^{(n)} \right) z_t^{(n)} \qquad (\text{since } p_t^{(0)} \equiv \log P_t^{(0)} = 0)$$

$$= \underbrace{-p_t^{(1)} z_t^{(1)} + \sum_{n=2}^{N} \left[\bar{a}_{n-1} + \bar{\mathbf{b}}_{n-1}' (\mathbf{c} + \mathbf{\Phi} \mathbf{f}_t) - \bar{a}_n - \bar{\mathbf{b}}_n' \mathbf{f}_t \right] z_t^{(n)}}_{\equiv B_t} + \left(\sum_{n=2}^{N} \bar{\mathbf{b}}_{n-1}' z_t^{(n)} \right) \boldsymbol{\varepsilon}_{t+1}$$

$$\text{(by (A1.8))}$$

$$= B_t + \left(\sum_{n=2}^{N} \bar{\mathbf{b}}_{n-1}' z_t^{(n)} \right) \boldsymbol{\varepsilon}_{t+1} \qquad (\text{by the definition of } B_t \text{ right above})$$

$$= B_t + \mathbf{d}_t' \boldsymbol{\varepsilon}_{t+1} \qquad (\text{by the definition of } \mathbf{d}_t \text{ in (A1.10)}). \tag{A1.12}$$

Since B_t defined here is known in date t, the source of the date t conditional variance of R_{t+1} is the last term, $\mathbf{d}_t' \boldsymbol{\varepsilon}_{t+1}$.

Appendix B. Derivation of (3.13)

Define $\mathrm{rp}_t^{(n)}$ as the left hand side of (A1.2). That is,

$$\mathrm{rp}_t^{(n)} \equiv \mathrm{E}_t \left(\log P_{t+1}^{(n-1)} \right) - \log P_t^{(n)} - y_t^{(1)} + \frac{1}{2} \bar{\mathbf{b}}_{n-1}' \mathbf{\Omega} \bar{\mathbf{b}}_{n-1}.$$

Replace t by $t + k$ and n by $n - k$, and take the expectation of both sides conditional on t, to obtain

$$\mathrm{E}_t \left(\mathrm{rp}_{t+k}^{(n-k)} \right) = \mathrm{E}_t \left(\log P_{t+k+1}^{(n-k-1)} \right) - \mathrm{E}_t \left(\log P_{t+k}^{(n-k)} \right) - \mathrm{E}_t \left(y_{t+k}^{(1)} \right) + \frac{1}{2} \bar{\mathbf{b}}_{n-k-1}' \mathbf{\Omega} \bar{\mathbf{b}}_{n-k-1}.$$

Sum this over $k = 0, 1, 2, \ldots, n - 2$ to obtain

$$\sum_{k=0}^{n-2} \mathrm{E}_t \left(\mathrm{rp}_{t+k}^{(n-k)} \right) = \mathrm{E}_r \left(\log P_{t+n-1}^{(1)} \right) - \log P_t^{(n)} - \sum_{k=0}^{n-2} \mathrm{E}_t \left(y_{t+k}^{(1)} \right) + \frac{1}{2} \sum_{k=0}^{n-2} \bar{\mathbf{b}}_{n-k-1}' \mathbf{\Omega} \bar{\mathbf{b}}_{n-k-1}.$$

(3.13) obtains because $\log P_t^{(n)} = -n y_t^{(n)}$ and $\log P_{t+n-1}^{(1)} = -y_{t+n-1}^{(1)}$. A similar formula can be found in Cochrane and Piazzesi (2008, Eq. (6)). Their formula does not involve the Jensen inequality term $\frac{1}{2} \bar{\mathbf{b}}_{n-k-1}' \mathbf{\Omega} \bar{\mathbf{b}}_{n-k-1}$ because their definition of the risk premium is the expected value of excess log holding-period return $\log P_{t+1}^{(n-1)} - \log P_t^{(n)} - y_t^{(1)}$.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rie.2018.04.003.

References

Bernanke, B., 2010. The economic outlook and monetary policy", speech delivered at the federal reserve bank of kansas city economic symposium. Jackson Hole, Wyoming, August 27.

Cochrane, J., Piazzesi, M., 2008. Decomposing the yield curve. Mimeo, March.

D'Amico, S., King, T., 2013. Flow and stock effects of large-scale treasury purchases" evidence on the importance of local supply. J. Financ. Econ. 108 (May(2)), 425–448.

Greenwood, R., Hanson, S., Vayanos, D., 2015. Forward guidance in the yield curve: Short rates versus bond supply. NBER Working Paper No. 21750, December 2015.

Greenwood, R., Vayanos, D., 2014. Bond supply and excess bond returns. Rev. Financ. Stud. 27 (3), 663–713.

Gurkaynak, R.S., Sack, B., Jonathan, H.W., 2007. The u.s. treasury yield curve: 1961 to the present. J. Monet. Econ. 54 (November(8)), 2291–2304.

Hamilton, J.D., Wu, J.C., 2012. The effectiveness of alternative monetary policy tools in a zero lower bound environment. J. Money Credit Banking 44 (February(1)), S3–S46.

Li, C., Wei, M., 2013. Term structure modeling with supply factors and the federal reserves large-scale asset purchase programs. International Journal of Central Banking 9 (March(1)), 3–40.

McCafferty, S., Driskill, R., 1980. Problems of existence and uniqueness in non-linear rational expectations models. Econometrica 48 (July(5)), 1313–1317.

Modigliani, F., Sutch, R., 1966. Innovations in interest rate policy. American Economic Review 56 (March(1/2)), 178–197.

Poloni, F., 2013. Quadratic vector equations. Linear Algebra Appl. 438, 1627–1644.

Tobin, J., 1969. A general equilibrium approach to monetary theory. J. Money Credit Banking 11 (February(1)), 15–29.

Vayanos, D., Vila, J.-L., 2009. A preferred-habitat model of the term structure of interest rates. NBER Working Paper No. 15487.