Introduction and summary

The secondary U.S. Treasury market is among the largest, most liquid, and most important financial markets worldwide. Daily trading volume in 2011 averaged $567.8 billion, more than tenfold the volume at the New York Stock Exchange (NYSE). The market is open around the clock, with trading involving both U.S. and international participants. Competition among dealers and brokers typically results in low bid–ask spreads, low brokerage fees, and fast order execution (for example, Fleming, 1997). Such features make the market very liquid across a wide spectrum of maturities.

Arbitrage is the practice of taking advantage of a price differential between securities that pay out similar cash flows (we provide a more rigorous definition at the beginning of the next section). This concept has immediate application in the U.S. Treasury market. For instance, consider two alternative investment strategies. The first entails purchasing a ten-year Treasury note. The second involves an investment of the same amount in a three-month Treasury bill that we repeatedly roll over at maturity into a newly issued three-month bill. For markets to clear, and absent market frictions, the price of the ten-year note needs to reflect investors’ expectations about the future path of the three-month Treasury rate during the next ten years. These expectations involve an adjustment to compensate risk-averse investors for bearing the risk that the price of the ten-year note will fluctuate during the holding period. If Treasury yields were to violate this condition, in a well-functioning capital market arbitrage trading would move funds across assets until prices adjust to balance out profit opportunities. By the same argument, yields on Treasury securities with various maturities will satisfy similar cross-sectional restrictions.

The Federal Reserve exploits the linkage across the term structure of bond yields to influence the availability and cost of money and credit in the economy. For instance, the Federal Open Market Committee (FOMC) uses open market operations to achieve a desired target rate in the federal funds market, where depository institutions lend balances at the Federal Reserve to other depository institutions overnight. Changes in the federal funds rate trigger a chain of events that affect other short-term interest rates.
foreign exchange rates, and the amount of money and credit. Most people, however, care especially about the cost of long-term credit—many firms rely on long-term debt to fund capital investment, and households take on long-term loans to buy their homes and cars. These observations underscore the importance of term structure models that help us gauge the effect of monetary policy actions (which typically impact the short end of the term structure) on long-term yields and, ultimately, a range of economic variables, including employment, output, and the prices of goods and services.

In this article, we discuss the pricing of U.S. Treasury securities via no-arbitrage arguments. We initially define what an arbitrage is and provide an intuitive one-period example that shows how to construct an arbitrage investment strategy in a frictionless capital market. We argue that absent transaction costs, information asymmetries, and other market imperfections, investors will trade away arbitrage opportunities. This will discipline the movement in prices of assets that are exposed to the same source of risk. We then formalize this intuition in the classical no-arbitrage term structure model of Vasicek (1977). We show that no-arbitrage arguments restrict the amount of return that investors demand in compensation for bearing a unit of risk (the so-called market price of risk) to be identical across the cross section of bonds. Exploiting this condition, Vasicek obtains a bond pricing formula that expresses the price of bonds of various maturities as a function of the spot interest rate, the market price of risk, and other model parameters.

This discussion also highlights the limitations of the Vasicek model. First, Vasicek assumes the market price of risk to be exogenous—his approach is silent about the economic forces that determine the amount of compensation investors require to bear risk. To clarify this link, we recast his model in a general equilibrium setting. This analysis shows that the market price of risk depends in fact on economic fundamentals such as the investors’ attitude toward risk and the volatility of the growth rate in aggregate consumption.

Second, in the Vasicek model a single variable, the spot interest rate, explains the fluctuations in the entire cross section of Treasury yields. One implication of this assumption is that bond yields and their changes are perfectly correlated. Correlations in pairs of yields with different maturities are positive and high in the data; however, they decrease considerably as the time to maturity of bonds becomes further apart. This feature suggests that additional factors might drive the U.S. Treasury yield curve and motivates a vast literature that extends the class of no-arbitrage term structure models to include multiple factors. We present an overview of this class of models, with an emphasis on the specifications that, similar to Vasicek’s model, allow for tractable bond pricing formulas (the so-called affine dynamic term structure models).

Third, the predictions of no-arbitrage models hinge on the critical assumption that markets are “perfect.” In order to take advantage of arbitrage opportunities, investors require access to capital. To trade away price misalignments, they need to be able to exchange securities at minimal cost based on information that is available to, and readily interpretable by, all investors. Clearly, no market satisfies all these conditions, and frictions typically become more severe during times of market stress. In extreme cases, markets could become segmented and arbitrage opportunities remain unexploited because of balance-sheet capacity limitations or because of higher-than-normal uncertainty and risk aversion. These conditions could reduce the effectiveness of no-arbitrage pricing arguments, possibly to a point where prices deviate from fundamental values.

Most of the time, frictions in the U.S. Treasury market are small. For instance, bid–ask spreads and other transaction costs are usually very low, and investors can trade securities with ease (for example, Fleming, 1997). Financial and economic crises typically do not impair these conditions. In fact, a flight to quality and/or liquidity can increase the demand for U.S. government debt, especially the most recently issued short-maturity nominal Treasury securities. This happened, in particular, during the recent financial crisis, when investors displayed a desire to hold only the safest and most liquid assets (for example, Gorton and Metrick, 2011; and Krishnamurthy, 2010). Nonetheless, government debt markets can exhibit some degree of segmentation because of the preferences of some investor clienteles (for example, pension funds, insurance companies, and other institutional investors) to hold securities that have specific maturities. So-called preferred habitat theories argue that these preferences could limit the substitutability of short- and long-term Treasury securities, distorting their relative pricing; capital constraints and risk aversion might prevent arbitrageurs from eliminating such profit opportunities. In the last part of the article, we expand on this discussion, focusing on the literature that studies limits to arbitrage in the government debt market.

Fourth, the dynamic term structure models that we review here typically rely on latent factors (or linear combinations of yields) to explain the variation in Treasury yields. Thus, this framework does not explain how bond yields respond to macroeconomic shocks, as these factors are void of immediate economic interpretation. Similarly, these models are silent about
the effect of monetary policy on economic variables, such as unemployment, gross domestic product (GDP) growth, and consumer prices. In response to these shortcomings, several recent studies explore the linkage between U.S. Treasury securities and the macroeconomy in no-arbitrage term structure models. We touch upon these issues at the very end, and postpone further discussion to the future.

No-arbitrage pricing in a one-period example

An arbitrage is an investment strategy that entails a nonpositive initial cost to generate a nonnegative cash flow that is positive with positive probability at some future date. Arbitrage opportunities should not exist in a frictionless market. Without transaction costs, information asymmetries, and other market imperfections, investors would immediately take advantage of any arbitrage opportunity. By doing so, they will close any misalignment in prices: Excess demand will push up the cost of securities that are relatively undervalued, and excess supply will lower the price of overvalued assets. Thus, no-arbitrage trading guarantees that securities are priced to reflect their future cash flow stream.

As a simple illustration of this concept, consider the case of an investor who trades in two assets at prices $P_t^i(t)$ and $P_t^{i+1}(t)$ on date $t$. The two securities do not pay dividends and are exposed to the same source of risk, so that their returns from $t$ to $t+1$ are described by the model

1) \[ \frac{\Delta P_t^i(t)}{P_t^i(t)} = \frac{P_t^i(t+1)/P_t^i(t)}{P_t^{i+1}(t)} = \mu_i + \Sigma_i \varepsilon(t+1), \quad i = 1, 2. \]

Here, $\mu_i$ denotes the constant expected rate of return on security $i$ during the unit interval, while the stochastic term $(\Sigma_i \varepsilon)$ is a mean zero innovation in the rate of return, with constant variance $\Sigma_i^2$. Being subject to the same shock $\epsilon$, the returns on the two assets by construction are perfectly correlated. Thus, the investor can exploit the co-movement in the two securities to eliminate risk from her portfolio. Suppose that she sells short $W_1$ worth of the first security’s shares and places a wealth amount $W_2$ in the second security. At time $t$, the portfolio is worth $W = W_2 - W_1$ in wealth. During the interval from $t$ to $t+1$, the change in wealth is determined by the rate of return on the two securities over that interval,

2) \[ \Delta W(t) = W_t(t+1) - W_t(t) = W_2(t) \frac{\Delta P_t^i(t)}{P_t^i(t)} - W_1(t) \frac{\Delta P_t^i(t)}{P_t^i(t)}. \]

Substituting the expression for the rate of return and rearranging the terms, we simplify equation 2 to

3) \[ \Delta W(t) = (W_2(t) - W_1(t)) \mu_i + (W_2(t) - W_1(t)) \Sigma_i \varepsilon(t+1). \]

An appropriate choice of $W_1$ and $W_2$ eliminates uncertainty in the strategy’s return. In particular, if the investor sets $W_1 = W_2 / (\Sigma_1 - \Sigma_2)$ and $W_2 = W_2 / (\Sigma_1 - \Sigma_2)$, the second term in equation 3 vanishes and the rate of return on invested wealth over the interval from $t$ to $t+1$ simplifies to

4) \[ \Delta W(t) = \frac{\mu_2 - \mu_1}{\Sigma_1 - \Sigma_2}. \]

At this point, we want to rule out arbitrage opportunities. To this end, we need to have the rate of return in equation 4 equal the risk-free rate, $r$, which we assume to be constant in this example. Thus, we have the following condition:

5) \[ \frac{\mu_2 - \mu_1}{\Sigma_1 - \Sigma_2} = r. \]

Rearranging terms in equation 5, we obtain the market clearing condition that links the expected return on the two securities, in excess of the risk-free rate, per unit of return standard deviation:

6) \[ \frac{\mu_2 - r}{\Sigma_2} = \frac{\mu_1 - r}{\Sigma_1}. \]

We denote the common value for this ratio with $\lambda$:

7) \[ \lambda = \frac{\mu_1 - r}{\Sigma_1}, \quad i = 1, 2. \]

The ratio $\lambda$ measures the market price of risk; that is, it quantifies the amount of return that investors demand in compensation for a unit of risk that they bear. To rule out arbitrage opportunities, we must have the coefficients $\mu_i$ and $\Sigma_i$ that determine the returns on securities $i = 1$ and $2$ in equation 1 satisfy the condition in equation 7. Intuitively, this restriction ties the price of the first security to that of the second security.
The coefficient \( \Sigma^2 \) of the spot rate determined by the assessment, at time \( t \), is a linear-plus-constant function of the spot rate \( r \), and the quadratic variation of the process is the constant \( \Sigma^2 \). These restrictions help us to obtain a closed-form bond pricing formula, which we derive next.

In the next section, we explain how these prices are tied together.

**The Vasicek model**

Here, we follow the Vasicek (1977) framework closely. We let the length of the time interval shrink to zero and recast the example from the previous section in continuous time. This simplifies the exposition considerably and clearly conveys the intuition for the results.

Assume that the spot risk-free rate, \( r \), in a frictionless market follows a mean-reverting diffusion process

\[
dr = \kappa(\theta - r)dt + \Sigma dZ,
\]

where \( Z \) is a standard Brownian motion. Equation 8 is a continuous-time analogue to the return process in equation 1. The left-hand side has the instantaneous change in the spot interest rate, \( dr = r(t + dt) - r(t) \). Similar to equation 1, the right-hand side of equation 8 is the sum of the expected change in \( r \), conditional on the realization of the time \( t \) spot rate, as well as a random shock. In particular, the term \( \kappa(\theta - r) \) describes the conditionally deterministic component of the spot rate evolution, with the coefficient \( \kappa > 0 \) controlling the speed of mean reversion of the process \( r \) toward its long-run mean \( \theta \). The Brownian shock \( dZ = Z(t + dt) - Z(t) \) has Gaussian distribution with mean zero and variance \( dt \). It takes place of the mean zero shock \( \epsilon \) over the discrete time interval from \( t \) to \( t + 1 \) in equation 1, where \( \text{Var}(\epsilon) = \Delta t = 1 \). The coefficient \( \Sigma^2 \) represents the constant instantaneous variance of the stochastic fluctuations of the spot rate. Equation 8 satisfies the affine restrictions of Duffie and Kan (1996); that is, the drift term \( \kappa(\theta - r) \) is a linear-plus-constant function of the spot rate \( r \), and the quadratic variation of the process is the constant \( \Sigma^2 \). These restrictions help us to obtain a closed-form bond pricing formula, which we derive next.

In the Vasicek model, the spot rate \( r \) summarizes the uncertainty in the economy. In particular, the time \( t \) price of a zero-coupon bond with maturity date \( T \) is determined by the assessment, at time \( t \), of the evolution of the spot rate \( r \), with \( t \leq s \leq T \). \( \text{Ito’s formula} \) gives then the dynamics for the bond price \( P_t = P(r_t, \tau) \), where \( \tau = T - t \):

\[
dP = \left( \frac{\partial P}{\partial t} + \frac{\partial P}{\partial r} \kappa(\theta - r) + \frac{1}{2} \frac{\partial^2 P}{\partial r^2} \Sigma^2 \right) dt + \frac{\partial P}{\partial r} \Sigma dZ
\]

\[
= P\mu(r, \tau)dt + P\sigma(r, \tau)dZ,
\]

where we have suppressed time \( t \) subscripts and defined

\[
\mu(r, \tau) = \frac{1}{P} \left( \frac{\partial P}{\partial t} + \frac{\partial P}{\partial r} \kappa(\theta - r) + \frac{1}{2} \frac{\partial^2 P}{\partial r^2} \Sigma^2 \right),
\]

\[
\sigma(r, \tau) = \frac{1}{P} \frac{\partial P}{\partial r} \Sigma.
\]

Following steps similar to those of the previous example, we consider an investor who sells short \( W_1 \) worth of the bond with maturity \( T_1 \) and who places wealth \( W_2 \) in the bond with maturity \( T_2 \). This strategy is worth \( W = W_1 - W_2 \), in wealth at time \( t \), which evolves according to

\[
dW = W_1 \frac{dP(r, \tau_1)}{P(r, \tau_1)} - W_2 \frac{dP(r, \tau_2)}{P(r, \tau_2)} = \mu_w dt + \sigma_w dz,
\]

where \( \mu_w = W_1 \mu(r, \tau_1) - W_2 \mu(r, \tau_2) \) and \( \sigma_w = W_1 \sigma(r, \tau_2) - W_2 \sigma(r, \tau_1) \). The investor can choose \( W_1 \) and \( W_2 \) to dynamically hedge her portfolio. In particular, setting

\[
W_1 = \frac{\sigma(r, \tau_2)W}{\sigma(r, \tau_1) - \sigma(r, \tau_2)},
\]

\[
W_2 = \frac{\sigma(r, \tau_1)W}{\sigma(r, \tau_1) - \sigma(r, \tau_2)},
\]

eliminates risk from her investment; that is, \( \sigma_w = 0 \) and the second term in the right-hand side of equation 11 vanishes. Thus, the position is insulated from the stochastic shock \( dZ \), and the instantaneous rate of return on invested wealth simplifies to

\[
dW = \frac{\sigma(r, \tau_1)\mu(r, \tau_1) - \sigma(r, \tau_2)\mu(r, \tau_2)}{\sigma(r, \tau_1) - \sigma(r, \tau_2)} dt.
\]

To avoid arbitrage opportunities, we need to have the growth rate in wealth to equal the risk-free rate,
Rearranging terms, we obtain a condition similar to equation 6:

\[ 15) \frac{\mu(r, \tau) - r}{\sigma(r, \tau)} = \frac{\mu(r, \tau) - r}{\sigma(r, \tau)}. \]

That is, the market price of risk \( \lambda \) is a function of the sole state variable of the economy, \( r \), and is independent of the bond time to maturity \( \tau \),

\[ 16) \lambda(r) = \frac{\mu(r, \tau) - r}{\sigma(r, \tau)}, \quad \forall \tau \geq 0. \]

To obtain a closed-form bond pricing formula, Vasicek assumes the market price of risk is constant; that is,

\[ 17) \lambda(r) = \lambda_w. \]

Substituting the expression for \( \mu(r, \tau) \) and \( \sigma(r, \tau) \) from equation 10 in equation 16 yields a partial differential equation for the bond price \( P \):

\[ 18) \frac{\partial P}{\partial t} + (\kappa(\theta - r) - \Sigma \lambda_w) \frac{\partial P}{\partial r} + \frac{1}{2} \Sigma \Sigma \frac{\partial^2 P}{\partial r^2} - rP = 0, \quad T \geq t, \]

with terminal condition \( P(r, \tau = 0) = 1 \). The solution to this equation is exponentially affine in the spot rate \( r \); that is, there are functions \( \overline{A}(\tau) \) and \( \overline{B}(\tau) \) of time to maturity \( \tau \) such that

\[ 19) P(r, \tau) = \exp\{\overline{A}(\tau) + \overline{B}(\tau) r\}. \]

Thus, we obtain a closed-form expression for the term structure of interest rates. In particular, the yield \( y \) on the bond with maturity date \( T \) is affine in the spot rate \( r \):

\[ 20) y(r, \tau) = A(\tau) + B(\tau) r, \]

where \( A(\tau) = -\overline{A}(\tau) / \tau \) and \( B(\tau) = -\overline{B}(\tau) / \tau \).

**The determinants of the market price of risk**

The Vasicek (1977) bond pricing formula hinges on the principle that absent arbitrage opportunities, the return on a locally risk-free portfolio of bonds must equal the risk-free rate. This approach is silent about the sources of the market price of risk \( \lambda \), and it takes the spot risk-free rate dynamics in equation 8 as given. Here, we show that the Vasicek bond pricing formula is consistent with the solution of the intertemporal consumption decision problem of a representative investor. While we arrive at the same pricing formula, this general equilibrium approach restricts the properties of the market price of risk and the instantaneous risk-free rate \( r \), which become functions of the investor’s attitude toward risk and the parameters that govern the aggregate dividend process. These results are well known in the literature (for example, Cox, Ingersoll, and Ross, 1985). The discussion in this section follows Goldstein and Zapatero (1996) and Cochrane (2005) closely.

Consider a security with ex-dividend price \( p \) that represents a claim to the aggregate output of the economy, which is paid out to the holder of the security in the form of a dividend \( D \). We assume that the security generates an ex-dividend return

\[ 21) \frac{dp}{p} = \mu dt + \sigma dZ, \]

where \( \mu \) is the ex-dividend expected rate of return on security \( p \), \( \sigma^2 \) is the instantaneous variance of the stochastic fluctuation in security \( p \)'s return, and \( Z \) is a standard Brownian motion. The quantities \( \mu \) and \( \sigma \) are endogenous to the model and will be determined in equilibrium. In contrast, the aggregate dividend is exogenously given by

\[ 22) \frac{dD}{D} = \alpha dt + \xi dZ, \]

\[ d\alpha = \kappa(\overline{\alpha} - \alpha) dt + \nu dZ, \]

Consider now an infinitely lived representative investor who trades in the security \( p \) and maximizes her lifetime utility of consumption,

\[ 23) U(c_{s \geq t}) = E_t \left[ \int_t^\infty e^{\delta(s-t)} u(c_s) ds \right]. \]

Cochrane (2005) shows that the first-order condition for this problem generates the basic pricing equation,
24) \( p_cu'(c_t) = E_t \int_0^\infty e^{-\delta s} u'(c_{t+s}) D_{t+s} ds \),

which equates the marginal cost of acquiring the security today at price \( p_t \) to the marginal benefit generated by its future dividend stream. Defining the discount factor as \( \Lambda_t = e^{-\rho t} u'(c_t) \), we can rewrite equation 24 as:

25) \( p_t \Lambda_t = E_t \int_0^\infty \Lambda_{t+s} D_{t+s} ds \).

Consider now a strategy that buys security \( p_t \) at time \( t \) and sells it at time \( t + \Delta \). Equation 25 then yields

26) \( p_t \Lambda_t = E_t \int_0^\infty \Lambda_{t+s} D_{t+s} ds + E_t[\Lambda_{t+s} p_{t+s}] \).

For small \( \Delta \to 0 \), this can be approximated by:

27) \( 0 = \Lambda_t D_t dt + E_t[d(\Lambda_t p_t)] \).

Itô’s lemma yields \( d(\Lambda_t p_t) = p_t d\Lambda_t + \Lambda_t dp_t + dp_t d\Lambda_t \), so that equation 27 becomes:

28) \( 0 = \frac{D_t}{p_t} dt + E_t \left[ \frac{d\Lambda_t}{\Lambda_t} \frac{dp_t}{p_t} + \frac{d\Lambda_t}{\Lambda_t} \frac{dp_t}{p_t} \right] \).

Using equation 25 to price the (instantaneous) risk-free zero-coupon bond, we obtain an expression for the spot risk-free rate,

29) \( r_t dt = -E_t \left( \frac{d\Lambda_t}{\Lambda_t} \right) \).

Then, rearranging equation 28, we obtain an equilibrium condition for the expected rate of return on security \( p_t \), \( \mu_t \);

30) \( E_t \left[ \frac{dp_t}{p_t} \right] + \frac{D_t}{p_t} dt - r_t dt = -E_t \left[ \frac{d\Lambda_t}{\Lambda_t} \frac{dp_t}{p_t} \right] \).

Assume now that the investor has the power utility function \( u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma} \) with the coefficient of risk aversion \( \gamma \). By the definition of \( \Lambda_t \), the stochastic discount factor dynamics are

31) \( \frac{d\Lambda_t}{\Lambda_t} = -\delta dt - \gamma \frac{dc_t}{c_t} + \frac{1}{2} \gamma(1+\gamma) \left( \frac{dc_t}{c_t} \right)^2 \),

so that equation 30 becomes:

32) \( E_t \left[ \frac{dp_t}{p_t} \right] + \frac{D_t}{p_t} dt - r_t dt = \gamma E_t \left[ \frac{dc_t}{c_t} \frac{dp_t}{p_t} \right] \).

Equation 32 says that the expected excess cum-dividend return on security \( p_t \) is proportional to the risk aversion coefficient \( \gamma \). Thus, more-risk-averse investors demand a higher risk premium to hold \( p_t \). Moreover, the risk premium on \( p_t \) depends on the correlation between aggregate consumption growth and the return on \( p_t \), \( E_t \left[ \frac{dc_t}{c_t} \frac{dp_t}{p_t} \right] \). Thus, an investor will require a positive risk premium to hold a security that generates a high return when consumption growth is high, that is, when \( E_t \left[ \frac{dc_t}{c_t} \frac{dp_t}{p_t} \right] > 0 \). This is intuitive, as such security generates, in expectation, a low payoff when consumption is low. This property makes the security less valuable to the investor, who is risk averse and wishes to smooth her consumption profile.

Note that in equilibrium, aggregate consumption equals the aggregate dividend, and thus it has dynamics identical to those given in equation 22. Substituting the endowment growth rate in equation 29 yields an expression for the equilibrium risk-free rate:

33) \( r_t = \delta + \alpha \gamma - \frac{1}{2} \gamma(1+\gamma) \xi_t^2 \).

Itô’s lemma gives us the spot rate dynamics

34) \( dr_t = \kappa(\theta - r_t) dt + \Sigma dZ_t \),

where we have defined the coefficients \( \theta = \gamma \alpha + \delta - \frac{1}{2} \gamma(1+\gamma) \xi_t^2 \) and \( \Sigma = \gamma \). Equation 34 is identical to the spot rate dynamics in equation 8, as in the Vasicek (1977) model. However, via equilibrium arguments we have established a linkage between the coefficients \( \kappa, \theta, \Sigma \) and economic fundamentals (that is, the coefficients \( \kappa, \alpha, \delta, \xi, \) and \( \nu \) that govern the endowment dynamics in equation 22 and the risk aversion parameter \( \gamma \)).

To obtain a formula for the price \( P_t \) of a zero-coupon bond with maturity date \( T \), it is useful to compute the spot rate dynamics under the risk-adjusted probability measure \( Q \) (Harrison and Kreps, 1979). With the help of equation 32, we obtain:

35) \( dr_t = \kappa(\theta^0 - r_t) dt + \Sigma dZ_t^0 \),

where we have defined \( \theta^0 = \theta - \frac{\nu}{\xi} \) and \( dZ_t^0 = dZ_t + \gamma \xi_t dt \). Then, we have
where the conditional expectation $E^Q_t[\cdot]$ is computed under the risk-adjusted measure $Q$.

The spot rate in equation 35 is a continuous Markov process. Thus, the evolution of $r_t$ over the interval $(t, T)$, given the history up to time $t$, depends only on $r_t$. Equation 36 then implies that the bond price is a function of $r_t$, $P(t, r_t, T)$, and by Itô’s lemma we obtain:

$$37) \quad dP = \left[ \frac{\partial P}{\partial t} + \frac{\partial P}{\partial r} (\kappa(0-r) + \frac{1}{2}\frac{\partial^2 P}{\partial r^2} \Sigma^2) \right] dt + \frac{\partial P}{\partial r} \Sigma dZ.$$ 

Moreover, we can apply equation 32 to determine the expected rate of return on the zero-coupon bond, in excess of the spot rate:

$$38) \quad E_t \left[ \frac{dP}{P} - r dt \right] = \gamma E_t \left[ \frac{dc}{c} \frac{dP}{P} \right].$$

Combining equations 37 and 38, we derive the fundamental differential equation for bonds:

$$39) \quad \frac{\partial P}{\partial r} \left(\kappa(0-r) - \sum_{s=0}^T \gamma \xi_s \right) + \frac{1}{2} \frac{\partial^2 P}{\partial r^2} \Sigma^2 + \frac{\partial P}{\partial t} - rP = 0,$$

where $\xi$ is the diffusion coefficient of the aggregate endowment process given in equation 22. Equation 39 is identical to the partial differential equation of the Vasicek model (equation 18) with the restriction $\lambda_n = \gamma \xi$. Consequently, assumptions about investors’ preferences and their endowment pin down the specification of the market price of risk. Specifically, $\lambda_n$ is higher when the investor is more risk averse, $\gamma \uparrow$, and when consumption growth is more volatile, $\xi \uparrow$.

**Multifactor dynamic term structure models**

In the Vasicek (1977) model, a single factor, the spot rate $r$, explains the fluctuations in the entire term structure of interest rates. One implication of this assumption is that bond yields and their changes are perfectly correlated. A cursory glance at figure 1 shows that there are co-movements in yields with different maturities,
but such correlations are far from perfect. This is evident in Table 1, which reports pairwise correlations in yields and their changes, $corr(y_i, y_j)$ and $corr(\Delta y_i, \Delta y_j)$, for maturity pairs $\tau_i$ and $\tau_j$ ranging from one quarter to 20 years. While the correlations in both the monthly (panel A) and quarterly (panel B) series are positive, they decrease considerably as the time to maturity in the pairs of bonds becomes further apart. This feature suggests that additional factors might drive the term structure of U.S. Treasury yields.

The evidence in Table 2 lends additional support to this conclusion. It shows the percentage of the yields’ variation explained by the principal components (PCs) extracted from the panel of bond yields with one-, four-, 12-, 20-, 40-, and 80-quarter maturities. The first principal component has the highest explanatory power, accounting for more than 95 percent of the variation in monthly and quarterly yields. The second and third components account for virtually all of the residual variation in yields. This is well known in the term structure literature; for instance, Litterman and Scheinkman (1991)
show that the variation in U.S. Treasury rates is best captured by three factors, interpreted as changes in “level,” “slope,” and “curvature” of the yield curve.

Figure 2 clarifies this interpretation. The yields’ PCs are an orthogonal linear transformation of the yields’ series; they are constructed so that each component explains the highest fraction of residual variance in the original series and is orthogonal to the preceding PCs. Figure 2 shows the coefficients in the vector $B_{\tau}$ that multiply the yields to form the first three principal components, PC$^j$, $j = 1, 2, \text{ and } 3$, as a function of the yields’ maturity $\tau$. The coefficients associated with the first PC are roughly the same across the yields’ maturities. This suggests that PC$^1$ is a proxy for a level factor, that is, shocks to that factor result in a parallel shift in yields across maturities. Consistent with this view, the correlation between PC$^1$ and $y_{\tau}$, $\tau \in \{1, 4, 12, 20, 40, \text{ and } 80 \text{ quarters} \}$ ranges from 93.6 to 99.7 percent in monthly data; we find similar values in the quarterly series. This is also evident in figure 3, which shows that the pattern in PC$^1$ resembles the shape of the yields in figure 1.

In contrast, the coefficients of the second PC are increasing in yields’ maturity $\tau$, while those of the third one are U-shaped, as shown in figure 2. Thus, as in Litterman and Scheinkman (1991), PC$^2$ is a proxy for a slope factor (positive shocks to this factor are associated with lower short-maturity yields and higher long-maturity yields), while PC$^3$ is a proxy for curvature. Indeed, the correlation between PC$^2$ and a measure of the term structure slope, $(y_{80Q} - y_{1Q})$, exceeds 90 percent, and the correlation of PC$^3$ with a measure of curvature, $(y_{80Q} - 2y_{40Q} + y_{1Q})$, is higher than 83 percent.

Taken together, these empirical observations motivate a vast literature that extends the no-arbitrage term structure model class to include multiple factors. As in the Vasicek (1977) model, the no-arbitrage conditions restrict the relative pricing of bonds with different maturities while remaining silent about all other conditions that characterize the equilibrium in the economy. Consistent with the evidence that level, slope, and curvature factors capture virtually all variation in Treasury yields, much of this literature has focused on three-factor models.

To maintain tractability, most studies rely on so-called affine models. In line with Duffie and Kan (1996), Dai and Singleton (2000, 2003), and Piazzesi (2010), the short-term interest rate, $r(t)$, is an affine (that is, linear-plus-constant) function of a vector of state variables, $X(t) = \{x_i(t), i = 1, \ldots, N\}$:

\[
40) \quad r(t) = \delta_0 + \sum_{i=1}^{N} \delta_i x_i(t) = \delta_0 + \delta^*_X X(t),
\]

where the state vector $X$ evolves according to

\[
41) \quad dX(t) = \kappa(\Theta - X(t))dt + \Sigma \sqrt{S(t)}dZ(t).
\]
Equation 41 extends the state dynamics in the Vasicek (1977) model (equation 8) to include N latent factors. The $N \times N$ matrix $\kappa$ in the first term on the right-hand side of equation 41 captures the dependence of infinitesimal changes in each $x_i(t)$ variable on the state vector $X(t)$. Similar to equation 8, the state vector $X(t)$ reverts to its mean $\Theta$, which is now an $N$-dimensional vector of constants. The process $Z$ is an $N$-dimensional Brownian motion. However, unlike the Vasicek (1977) model, the instantaneous variance of the fluctuations in $X$ is no longer constant. It depends on the level of $X$ via the $N \times N$ diagonal matrix $S(t)$, which has $i$th diagonal element $s_{ii}(t) = \alpha_i + \beta_i^2 X_i(t)$.

To price bonds, we specify the market price of risk, $\Lambda(t)$. This is often assumed to depend on the state vector $X(t)$, rather than being constant, as in equation 17. For instance, Dai and Singleton (2000) set

$$\Lambda(t) = \sqrt{S(t)} \lambda,$$

where $\lambda$ is an $N \times 1$ vector of constants. This functional form guarantees that risk compensation goes to zero as the variance of the state vector vanishes—a condition that rules out arbitrage opportunities. However, Duffie (2002) notes that since the variance term is nonnegative, this structure limits the variability of the compensation that investors expect to receive for facing a given risk. In particular, he shows that this condition is restrictive as it prevents risk compensation to switch sign over time—a feature that is important to explain the variation in Treasury returns. He goes on to extend the market price of risk in a way that relaxes this restriction; subsequently, Duarte (2004) and Cheridito, Filipović, and Kimmel (2007) offer further generalizations.

Within this setting, the time $t$ price of a zero-coupon bond with time to maturity $\tau$ is given by

$$P(X, \tau) = \exp\{\tilde{A}(\tau) + \tilde{B}(\tau)'X(t)\},$$

where the functions $\tilde{A}(\tau)$ and $\tilde{B}(\tau)$ solve a system of ordinary differential equations (ODEs); see, for example, Duffie and Kan (1996). Thus, the yield $y$ on the bond with time to maturity $\tau$ is affine in the state vector $X$:

$$y(X, \tau) = A(\tau) + B(\tau)'X,$$

where $A(\tau) = -\tilde{A}(\tau)/\tau$ and $B(\tau) = -\tilde{B}(\tau)/\tau$. This is similar to equations 19 and 20 for the Vasicek (1977) model, except that the $N$-dimensional state vector $X$.
limits to arbitrage in the market of government debt

The models we present in this article hinge on the assumption that whenever an arbitrage opportunity arises, investors implement trading strategies to profit from it until asset prices change to drive risk-adjusted net expected returns to zero. In practice, however, prices might not converge if markets are not perfect. For instance, frictions such as transaction costs, leverage constraints, and limited availability of capital could hinder investors’ ability to trade away arbitrage opportunities. In this section, we first provide evidence that transaction costs in the U.S. Treasury market are small. We then explore the role of leverage and capital constraints in arbitrage trading. In particular, we argue that financial institutions relax these constraints by participating in a vast repo market in which U.S. Treasury securities are a valuable form of collateral. Next, we report some well-documented patterns in Treasury securities’ yields that can arise because of institutional constraints, arbitrage capital requirements, and market segmentation. We conclude by briefly considering the relevance of Treasury market frictions for monetary policy interventions during the recent financial crisis and for the specification and estimation of no-arbitrage term structure models.

Transaction costs and liquidity in the U.S. Treasury market

As we mentioned earlier, the secondary U.S. Treasury market is one of the largest and most important financial markets worldwide. The around-the-clock trading activity in this market, by both U.S. and international participants, far exceeds that observed on many popular exchanges.

While high trading volume is often used as an indicator of asset marketability, there is evidence that it could be a noisy, and possibly even poor, liquidity measure. Fleming (2003) shows that trading volume in the secondary U.S. Treasury market, as well as yields’ volatility, often peak during periods of market stress, when trading is more difficult than usual. In contrast, the difference between bid and ask Treasury prices (the so-called bid–ask spread) is a simple and more robust indicator of the ease with which investors can exchange securities. For instance, Fleming (2003) shows that bid–ask spreads on Treasury securities correlate more highly with popular liquidity indicators, such as price impact, defined as the sensitivity of price changes to net trading activity (the difference between buyer- and seller-initiated trades). Moreover, the bid–ask spread has an intuitive interpretation in terms of transaction costs that an investor would incur if she were to buy/sell securities. For these reasons, we focus on this measure of liquidity here.


Sources: Authors’ calculations based on intraday quotes data from GovPX; and Board of Governors of the Federal Reserve System staff’s calculations based on daily data from BrokerTec.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>10th</td>
</tr>
<tr>
<td><strong>A. Summary statistics for period June 17, 1991–June 15, 2001</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treasury bills’ bank discount rate bid–ask spreads</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three months</td>
<td>0.75</td>
<td>0.90</td>
<td>0/2</td>
</tr>
<tr>
<td>Six months</td>
<td>0.80</td>
<td>0.83</td>
<td>0/2</td>
</tr>
<tr>
<td>One year</td>
<td>0.71</td>
<td>0.72</td>
<td>0/2</td>
</tr>
<tr>
<td>Treasury notes’ prices bid–ask spreads</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two years</td>
<td>0.26</td>
<td>0.18</td>
<td>0/4</td>
</tr>
<tr>
<td>Five years</td>
<td>0.38</td>
<td>0.26</td>
<td>0/2</td>
</tr>
<tr>
<td>Ten years</td>
<td>0.40</td>
<td>0.29</td>
<td>0/2</td>
</tr>
<tr>
<td>Treasury notes’ prices bid–ask spreads</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two years</td>
<td>0.36</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Five years</td>
<td>0.48</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Ten years</td>
<td>0.84</td>
<td>0.47</td>
<td></td>
</tr>
</tbody>
</table>


Sources: Authors’ calculations based on intraday quotes data from GovPX; and Board of Governors of the Federal Reserve System staff’s calculations based on daily data from BrokerTec.
June 15, 2001. It is evident that bid–ask spreads are small across bond tenors. For instance, the median spread on bills is one-half of a basis point. Spreads remain low even in the right tail of the distribution (for example, the 90th percentile is one and a half basis points). Among Treasury notes, the two-year security appears to be the most liquid, with a median spread of one-quarter of a 32nd of a percentage point of par. Transaction costs remain low on longer-maturity Treasury securities, with a typical bid–ask spread of one-half of a 32nd of a percentage point. Table 3, panel B shows similar diagnostics using a more recent sample of Treasury prices from January 1, 2001, through January 31, 2012. At various maturities, spreads fall in a range from 0.36 to 0.84 32nds of a percentage point, and standard deviations are small, too. Taken together, this evidence confirms that investors can typically trade Treasury securities with ease across the term structure. Those who seek to take advantage of misalignments in prices can do so at low transaction costs.

**Leverage constraints and the availability of arbitrage capital**

A liquid secondary market is not necessarily enough to guarantee that Treasury prices will converge to their no-arbitrage equilibrium values. For instance, Gromb and Vayanos (2010) suggest that transaction costs are only one of the financial market inefficiencies that can pose limits to arbitrage. In a simple theoretical framework, they show that no-arbitrage pricing does not hold in asset markets when arbitrageurs face leverage constraints (for example, Gromb and Vayanos, 2002; Geanakoplos, 2003; and Gârleanu and Pedersen, 2011) as well as equity capital requirements (for example, Shleifer and Vishny, 1997). In this respect, the presence of a vast market for repurchase agreements (repos) facilitates arbitrage trading greatly. A repo is a transaction that combines a spot market sale with a simultaneous forward agreement to repurchase the underlying instrument at a later date, often the next day (for example, Duffie, 1996). Effectively, a repo is a collateralized loan. The loan amount equals the sale value of the security (typically given by the market price of the security minus a margin, the so-called haircut), while the repo rate is the interest on the loan. The counterparty in a repo contract, who provides the funds for the loan and earns interest at the repo rate, is said to engage in a reverse repo.

Access to the repo market provides financial institutions with arbitrage capital to finance their trading activity. For instance, if the price of an asset falls below its fundamentals, a dealer can purchase it in the secondary market. Concurrently, if the security constitutes an acceptable form of collateral, the dealer can pledge it in the repo market and thus obtain funds in the amount of the price of the security, net of the repo haircut. The funds borrowed against the security offset, up to the haircut, the cost to acquire it. Excess demand for the security will push its price up. If the price increase exceeds the cost of financing in the repo market, the dealer will reap a profit. Conversely, if a dealer perceives a security to be overpriced, the dealer can engage in a reverse repo. The dealer can then sell the (overpriced) collateral in anticipation that its price will fall. If that happens, the dealer will be able to buy the security back at a lower price on a later date, and use it to unwind the reverse repo.

Over the past decades, the repo market has grown dramatically in size and popularity (for example, Gorton and Metrick, 2011). On one side, mutual funds (especially money market funds), corporations, and state and local governments have been expanding their use of reverse repos to put their cash reserves to work while concurrently acquiring high-quality collateral for protection of their investment. On the other side, financial institutions have been increasingly relying on repos to finance their operations. For instance, figure 4 shows the outstanding value of repurchase and reverse repurchase agreements by primary dealers from 1996 through 2011. The outstanding value of repos on dealers’ books is very high, and it exceeds that of reverse repos. The increasing pattern in quantities is also evident, in spite of a large decline at the peak of the U.S. financial crisis in 2008–09. Yet, figure 4 greatly underrepresents the magnitude of the U.S. repo market, which is, in fact, imprecisely documented. Gorton and Metrick (2011) provide an overview of different sources that estimate it to be around $10 trillion in the late 2000s. These estimates include transactions taking place in the triparty repo market, in which clearing banks (JPMorgan Chase and the Bank of New York Mellon) provide clearing and settlement services to the lender (the cash investor) and the borrower (the collateral provider); see, for example, Copeland, Martin, and Walker (2011). Estimates by the Tri-Party Repo Infrastructure Reform Task Force at the Federal Reserve Bank of New York place the size of that market at nearly $1.7 trillion as of January 2012 (see table 4).

Treasury securities are a valuable form of collateral in repurchase agreements. Table 4 shows that they account for approximately a third of the notional value of the underlying securities in triparty repos (other categories include securities issued by corporations, federal agencies, and municipalities). Similar evidence holds in the bilateral repo market (for example, Copeland, Martin, and Walker, 2011). Moreover, when
pledged as collateral, U.S. government debt is subject to a haircut that is usually very small. The margin on short-term Treasury securities is typically around 2 percent. It is higher for longer-maturity bonds, which have a higher price sensitivity to interest rate fluctuations; nonetheless, at approximately 5–6 percent, it is below the margin on other securities that are forms of collateral in repurchase agreements. Such margins have been remarkably stable even during times of market stress. This is in stark contrast with haircuts on corporate bonds, asset-backed securities, and collateralized mortgage obligations that lacked the support of government guarantees.7

In sum, this discussion highlights that financial institutions can rely on a vast repo market to fund their arbitrage positions, especially in the Treasury market. This is evident from the sheer value of Treasury securities pledged as collateral in repo transactions. Moreover, small and stable haircuts on Treasury securities allow investors to finance a larger portion of their positions via repos, contributing further to relaxing capital and leverage constraints.

Yet, market frictions matter

Arbitrage opportunities across Treasury securities tend to disappear quickly as investors trade them away in a liquid secondary market, often using repos to finance their positions. Nonetheless, market frictions can still play an important role in this market.

The fact that newer vintages of Treasury bonds typically trade at a premium compared with older vintages is a classic example. This phenomenon is often documented by the spread between the yield for on-the-run bonds (the most recent issue of bonds with a certain maturity) and that for off-the-run bonds (older issues of bonds with the same tenor). This evidence is puzzling, as the cash flows associated with two long-run (for example, 30-year) bonds are similar, even though the bonds are issued six months apart. It motivates a convergence trade that involves the purchase of the (cheaper) off-the-run bond and a short position in the

<table>
<thead>
<tr>
<th>Asset group</th>
<th>Collateral value (billions of U.S. dollars)</th>
<th>Percentage of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. Treasury securities</td>
<td>567.31</td>
<td>34</td>
</tr>
<tr>
<td>Other</td>
<td>1,098.93</td>
<td>66</td>
</tr>
<tr>
<td>Total</td>
<td>1,666.24</td>
<td>100</td>
</tr>
</tbody>
</table>

Notes: The table summarizes the activity in the triparty repo market for different types of collateral as of January 11, 2012. The “other” category includes repos collateralized with corporate bonds, federal agencies’ securities, and municipality debt.


---

7. In sum, this discussion highlights that financial institutions can rely on a vast repo market to fund their arbitrage positions, especially in the Treasury market. This is evident from the sheer value of Treasury securities pledged as collateral in repo transactions.
(more expensive) on-the-run security. The spread between old vintages of bonds tends to narrow as time goes by; thus, absent market frictions, a convergence trade would generate an arbitrage profit. In practice, arbitrageurs attempting to trade this strategy engage in a reverse repo to establish a short position in the on-the-run bonds (see the previous subsection). Since these bonds are in limited supply, excess demand for this collateral pushes repo rates below the market interest rate. This creates a significant cost of carry associated with the convergence trade, which erodes profits (for example, Duffie, 1996; and Krishnamurthy, 2002). Thus, a positive spread between off-the-run and on-the-run bond yields is not an arbitrage as long as the spread in repo rates compensates for the yield differential. Yet, the puzzle remains: Why are new bonds more expensive than old ones? Duffie (1996) and Krishnamurthy (2002) note that this situation can arise when some investors have a preference for liquidity and are restricted from participating in the repo market. For example, fixed-income mutual funds tend to hold liquid on-the-run bonds (similar to those included in the bond indexes to which they benchmark their performance).

The market for Treasury Inflation-Protected Securities (TIPS) provides another striking example. The U.S. Department of the Treasury started to issue TIPS in 1997. In the early stages of TIPS life, secondary market liquidity was very limited and TIPS traded at a discount (for example, Ajello, Benzoni, and Chyruk, 2011; D’Amico, Kim, and Wei, 2010; Haubrich, Pennacchi, and Ritchken, 2010; and Pflueger and Viceira, 2011). By 2004, the liquidity premium in TIPS yields had declined considerably as trading became more active in the TIPS market. More recently, the TIPS market experienced new significant disruptions across the term structure, with the five-year TIPS rate climbing above 4 percent in fall 2008. Fleenkerstein, Longstaff, and Lustig (2010) go one step further and argue that TIPS prices allow for arbitrage opportunities. In particular, they suggest a strategy that involves buying TIPS and selling inflation protection in the inflation swap market. They fine-tune the position to replicate the cash flows of a nominal bond and conclude that TIPS are undervalued relative to nominal Treasury securities. The strategy, however, involves committing arbitrage capital for the duration of the investment (possibly a long period of time), with the risk that if liquidity conditions deteriorate and investors are forced to unwind the position, they might incur a loss. These concerns, combined with disruptions in the TIPS and inflation swap markets, might have contributed to pushing the price differential up, especially in the fall of 2008, during the financial crisis.

These examples suggest that investors’ demand for bonds could depend on factors that go beyond the maturity structure of the cash flow and the issuer’s default risk. In the next subsection, we expand on these ideas.

**Preferred habitat theories**

One relevant implication of the absence of arbitrage in the market for Treasury securities is a perfect degree of substitutability across bond maturities—investors are willing to absorb any amount of bonds at their equilibrium prices. Shocks to the net supply of, or demand for, bonds of one maturity do not affect other yields, nor the shape of the term structure of interest rates. Early empirical studies that tested this condition in the U.S. Treasury market could not identify violations of the no-arbitrage principle. In particular, several papers (for example, Modigliani and Sutch, 1966; and Ross, 1966) evaluate the effectiveness of the so-called Operation Twist. Between 1962 and 1964, the Federal Reserve and the U.S. Department of the Treasury started selling short-term government bonds while purchasing long-term ones. The policy objective was to flatten the slope of the term structure by raising short-term interest rates to improve the balance of payments while lowering long-term rates to stimulate private investment. None of the papers found a significant effect of Operation Twist on the level of yields across the term structure.

These results discouraged further attempts to explore early theories that introduced limits to arbitrage across Treasury securities of different maturities in the form of investors’ preferred habitat, demand/supply pressure, and bond market segmentation (for example, Culbertson, 1957; and Modigliani and Sutch, 1966). According to these theories, various classes of investors have well-defined preferences for specific maturities. Pension funds and life insurance companies, for example, purchase bonds of longer maturities, while banks buy short-term securities. Because of such differences in preferences or regulatory requirements, bonds of different maturities end up being imperfect substitutes. Consequently, equilibrium yields are determined by the interaction between the demand by various clienteles and the aggregate bond supply for each specific maturity.

More recently, new evidence has been supporting the view that there are limits to arbitrage in government bond markets, consistent with preferred habitat theories. For example, Greenwood and Vayanos (2010b) study the consequences of the Pensions Act 2004, which reformatted the UK pension system. The act introduced capital requirements to ensure the solvency of pension funds and anchored the evaluation of their liabilities to long-term interest rates. These institutional changes...
promoted pension funds to hedge their liabilities against interest rate risk and shifted their demand toward long-term government bonds. While these events were unfolding, there was a simultaneous drop in long-term yields. This evidence is not inconsistent with demand pressure and habitat preference theories, and it is difficult to explain based solely on the notion of sudden changes in either interest rate expectations or fundamental risk within the framework of a no-arbitrage model. Similarly, Greenwood and Vayanos (2010a, b) also document evidence of demand pressure in the U.S. Treasury market. Between March 2000 and December 2001, the U.S. Department of the Treasury repurchased 10 percent of the long-term government bonds outstanding as of December 1999. This intervention reduced the spread between the 20- and five-year yields by 65 basis points in a few weeks and contributed to the inverting of the term structure slope.

**Implications for monetary policy**

Recent interventions of the Federal Reserve in the government bond markets, known as large-scale asset purchases (LSAPs), have revived interest in the market segmentation hypothesis and its applications to the nominal yield curve. After a first round of LSAPs directed to the stabilization of the government agency bond market in late 2008 (known as “quantitative easing 1,” or “QE1”), the Federal Reserve started purchasing long-term Treasury bonds in 2009 and stepped up its demand with a second purchase program of $600 billion from November 2010 through June 2011 (often referred to as “quantitative easing 2,” or “QE2”).

Several recent empirical studies assess the effect of the Federal Reserve’s purchases of long-term Treasury securities and other bonds on interest rates (for example, D’Amico and King, 2010, D’Amico et al., 2012; Gagnon et al., 2010; Hancock and Passmore, 2011; and Krishnamurthy and Vissing-Jørgensen, 2010, 2011). This literature attempts not only to quantify the effect of LSAPs on different yields, but also to identify the channels through which these unconventional monetary policy interventions work.

A direct comparison of their findings is difficult because of differences in data, sample frequency, and approaches used to disentangle various channels. There is some agreement that LSAPs have been effective in lowering medium- and long-term rates. However, the channels through which this policy works are more controversial. For instance, Krishnamurthy and Vissing-Jørgensen (2011) find evidence for a signaling channel, a unique demand for long-term safe assets, and an inflation channel for both QE1 and QE2; and they find evidence for a mortgage-backed securities prepayment channel and a corporate bond default risk channel for QE1. They argue that Treasury-securities-only purchases in QE2 had a disproportionate effect on Treasury and agency securities relative to mortgage-backed and corporate securities, with yields on the latter falling primarily through the market’s anticipation of lower future federal funds rates. This is consistent with the view that QE2 constitutes a commitment by the Federal Reserve to keep interest rates low in the future: Lower expected future spot rates push long-term yields down regardless of market segmentation (Clouse et al., 2003; and Eggertsson and Woodford, 2003).

In contrast, D’Amico et al. (2012) and Gagnon et al. (2010) conclude that reductions in interest rates primarily reflect lower risk premiums rather than lower expectations of future short-term interest rates. This is consistent with the view that LSAPs reduce duration risk and create a scarcity effect on long-term bonds that are in high demand among some investor clienteles.

While empirically challenging, disentangling the relative importance of various channels is critical to guide monetary policy. On one side of the debate, the findings for QE2 by Krishnamurthy and Vissing-Jørgensen (2011) raise the question of whether the main impact of a Treasury-securities-only QE may have been achievable with a Federal Reserve statement committing to lower federal funds rates (a policy that does not require the Federal Reserve to commit its balance sheet). On the opposite side of the debate, the conclusions of D’Amico et al. (2012) and Gagnon et al. (2010) support the use of LSAPs in the Treasury market as a powerful tool of monetary policy easing when the federal funds rate is at the zero lower bound. Moreover, this debate has interesting implications for no-arbitrage term structure models, which we discuss in the next subsection.

**Implications for no-arbitrage term structure models**

Recent developments in the limits-to-arbitrage literature are useful to sharpen the specification of no-arbitrage term structure models. For instance, Krishnamurthy and Vissing-Jørgensen (2011) suggest that QE2 has affected long-term Treasury yields mainly by lowering the market’s expectations of future federal funds rates. To accommodate this evidence, one could extend the term structure models discussed in this article to allow for changes in the way agents form expectations about future spot rates. Models that allow for regime switches in monetary policy (for example, Ang et al., 2011; Bikbov and Chernov, 2008; and Fuhrer, 1996) and evolving beliefs about inflation dynamics (for example, Sargent, 2001) are a useful step in this direction. One challenge is to extend the analysis to an environment in which the federal funds rate is at the zero lower bound.
Moreover, preferred habitat theories motivate structural, theoretically founded restrictions on the dynamics of yields that could be useful to refine existing dynamic term structure models. This is an interesting area of research that has seen considerable progress in the past few years. Hamilton and Wu (2012), for example, follow Greenwood and Vayanos (2010a, b) and Doh (2010) in using the promising theoretical framework developed by Vayanos and Vila (2009) to rationalize and evaluate the Federal Reserve’s large-scale purchases of U.S. Treasury securities across different maturities. In their models, risk-averse arbitrageurs interact with preferred habitat investors, whose demand for a bond with a specific maturity is an increasing function of its yield. Hamilton and Wu (2012) introduce this market segmentation in an affine term structure model and conclude that the maturity structure of debt held by the public affects the level, slope, and curvature of the yield curve. In this setting, they find that bond demand shocks have a significant effect on bond prices, even in the presence of a binding zero lower bound constraint for the federal funds rate (see also related evidence in Krishnamurthy and Vissing-Jørgensen, 2010).

Finally, the liquidity differential often observed across bond vintages, which we discussed earlier, raises the question about which Treasury yield series are more suitable for the estimation of dynamic term structure models. Most empirical studies have been focusing on liquid on-the-run securities. However, some researchers have advocated using off-the-run bonds, which include a smaller liquidity premium compared with new issues (for example, Gürkaynak, Sack, and Wright, 2007). More broadly, this discussion highlights the challenge to choose an appropriate measure for the risk-free rate. To what extent is the ability to trade the security with ease a defining feature of the risk-free asset? In principle, one could explicitly model the liquidity wedge across yields to identify the “true” term structure of interest rates. This approach could be particularly useful when modeling segments of the Treasury market that are more sensitive to liquidity disruptions (for example, the TIPS market).11

**Conclusion**

In this article, we discuss the role of arbitrage trading in the U.S. Treasury market. We start out by defining the concept of arbitrage and illustrate it in a simple one-period example. We then show how the absence of arbitrage aligns risk-adjusted returns across bonds with different maturities in the framework of the Vasicek (1977) one-factor term structure model. Along the way, we explain the link between bond risk premiums and the underlying economy in a stylized general equilibrium setting. Empirical evidence on bond yields suggests that at least three factors drive fluctuations in the term structure of interests rates. This observation motivates a vast literature on multi-factor models, which we briefly review with an emphasis on tractable affine specifications. The article ends with an evaluation of market frictions in the government debt market and their implications for no-arbitrage term structure models.

In the models we discuss here, the factors are typically latent variables (or linear combinations of yields) void of immediate economic interpretation. Thus, these models are silent about the response of bond yields to macroeconomic shocks, as well as the chain of events through which monetary policy intervention ultimately impacts the real economy. Early studies investigate these questions by directly relating current bond yields to past yields and macroeconomic variables in a vector autoregression framework (for example, Estrella and Mishkin, 1997; and Evans and Marshall, 1998, 2007). More recently, much work has gone into incorporating macroeconomic information in no-arbitrage dynamic term structure models. We postpone further discussion of this literature to the future.
NOTES

1Authors’ calculations based on data from the Securities Industry and Financial Markets Association (SIFMA) and the New York Stock Exchange (NYSE) Facts and Figures (formerly the online NYSE Fact Book). The data are available at www.sifma.org/research/statistics.aspx and www.nydata.com/factbook. The label “secondary” market refers to the market in which Treasury bills, notes, and bonds are traded once they are issued. This label sets the market apart from the “primary” market in which these securities are first auctioned and sold by the U.S. Department of the Treasury.

2See the “About the FOMC” section at www.federalreserve.gov/monetarypolicy/fomc.htm.

3Treasury bill prices are quoted on a bank discount rate basis with tick size of 1 basis point. Treasury notes and bonds are quoted at percentage of par in 32nds of a point. See, for example, Sundaresan (2001) for more information on trading practices in the secondary U.S. Treasury market.

4The two-year note is the shortest-maturity coupon-bearing security issued by the U.S. Treasury. This makes it appealing to people who seek a medium-term investment that comes with the convenience of regular coupon payments.

5In the United States, retail depositors at a bank insured by the Federal Deposit Insurance Corporation (FDIC) are entitled to interest payments and are reimbursed up to a certain amount if the bank fails. Limits to the amount of deposit insurance reduce the appeal of demand deposit accounts to corporations. Under the Federal Reserve’s Regulation Q, as in effect until July 21, 2011, corporations were not entitled to earn interest on demand deposit accounts. In contrast, engaging in a reverse repo allows institutional investors to earn interest at lower risk because of the presence of collateral.

6Most estimates of the repo market size rely on surveys of its participants. Adrian et al. (2012) provide an overview of data requirements necessary to monitor repos and securities lending markets for the purposes of informing policymakers and researchers about firm-level and systemic risk. They conclude that data collection is currently incomplete, and argue that a comprehensive collection should include six characteristics of repo and securities lending trades at the firm level: principal amount, interest rate, collateral type, haircut, tenor, and counterparty.

7See, for example, Gorton and Metrick (2011) and Krishnamurthy (2010) for evidence based on the bilateral repo market. Margins and funding were mostly stable during and after the crisis period in the triparty repo market, except in rare cases when funding dropped precipitously (Copeland, Martin, and Walker, 2011).

8Convergence trades were important positions in the portfolio of the Long-Term Capital Management (LTCM) fund. These trades received considerable attention in the news in 1998, when an increase in the spread between off-the-run and on-the-run bond yields produced significant losses for LTCM. The fund was eventually liquidated. See, for example, Lewis (1999).

9Recently, Swanson (2011) revisits this episode using an event-study approach that matches high-frequency changes in financial markets within narrow windows of time around major, discrete announcements to measure the effects of those announcements. He finds some support for the notion that Operation Twist performed as its designers thought it would.

10This is not, however, a unanimous view. For a dissenting voice, see, for instance, Cochrane (2011), who states: “QE2 doesn’t seem to have lowered any interest rates. Yes, five-year rates trended down between announcements, though no faster than before. The November [2010] QE2 announcement and subsequent purchases coincided with a sharp Treasury rate rise. The five-year yields where the Fed bought most heavily didn’t decline relative to the other rates, as the Fed’s ‘segmented markets’ theory predicts. The corporate and mortgage rates that matter for the rest of the economy rose throughout the episode.”

11Recent work by D’Amico, Kim, and Wei (2010) is an interesting example.
REFERENCES


Ajello, Andrea, Luca Benzoni, and Olena Chyruk, 2011, “Core and ‘crust’: Consumer prices and the term structure of interest rates,” Federal Reserve Bank of Chicago, working paper.


