Chapter Fourteen

Recent Advances in Old Fixed-Income Topics: Liquidity, Learning, and the Lower Bound

Jean-Sébastien Fontaine and René Garcia
Bank of Canada and EDHEC Business School

14.1 Introduction

The expectations hypothesis (EH) paradigm dominates the world of term structure modeling. Formal tests of the hypothesis go back at least to Meiselman (1962), but the possibility that long-term rates provide biased forecasts of future short rates has long preoccupied economic discourse (see, e.g., the discussion in Culbertson, 1957). True, bond yields are the sum of the expected interest rate and a risk premium, but this blurs different types of deviations from the EH. What distinguishes modern literature is the emphasis on interest rate risk as the leading (or sole) determinant of the risk premium. This is apparent in, for instance, the reviews by Dai and Singleton (2003), Piazzesi (2005), Gurkaynak and Wright (2012), and Duffee (2013). The older literature emphasized several mechanisms that could cause deviations from the expectations hypothesis.

Sources of risk premium other than interest rate risk found a refuge in undergraduate textbooks while the academic agenda leapt forward, developing an array of sophisticated yet tractable no-arbitrage models. Tractability may lead to progress, nudging aside some old themes for a time, but the empirical evidence eventually catches up as theory matures. In this chapter, we selected old topics that have returned to the forefront in fixed-income research: liquidity, learning and the bound underlying nominal yields. Progress has been been rapid in these fields in recent years.

First, the importance of liquidity and of intermediation frictions has reemerged in the face of large deviations from no-arbitrage relationships during and around the 2008 financial crisis. The prices of neighbouring Treasury bonds became widely misaligned, to an extent unjustified by the small differences in their cash flows. But the fact that similar bonds often bear dissimilar prices had not escaped those estimating smooth zero-coupon curves from observed coupon bond prices (e.g., Bliss, 1987). This dispersion of prices has been attributed to bid-ask spreads, search frictions, short-selling costs and clientele effects, but was generally felt to be unimportant in broader contexts. We discuss recent advances, showing how these apparent deviations from arbitrage reveal important state variables and possibly generate their own risk premium.

Second, several recent works have revisited an old conundrum: testing whether yields conform to the EH relies on the auxiliary assumptions that investors' expectations are rational and based on a complete information set. Combining observed yields with data from surveys of yield forecasts can break this conundrum, and early results confirms that expectational "errors" could explain part of the evidence against the EH. Modern term structure models use a flexible reduced-form specification of the pricing kernel and should capture shifts in investors' forecasts within an encompassing risk premium. Nonetheless, the required variability of the pricing kernel appears hard to square with macro risk implicit in the evolution of consumption growth or the inflation rate. Incorporating the subjective expectations embodied in survey data as well as the effect of learning within term structure models, whether structural or based on a no-arbitrage argument, has produced a stimulating research agenda.

Our final topic builds on the observation that nominal rates are bounded from below. As the Federal Reserve target rate hovers near zero for yet another year—since the end of 2008— and as Japan, Europe and other countries reach the lower bound, the challenge to benchmark models continues to grow. In the previous era of relatively high rates, the literature had escaped from the limitations of the square-root model and taken refuge in Gaussian high grounds. The lower bound underlying yields changes most stylized facts about yield, requiring different but yet tractable models.

Section 14.2 first reviews how differences in prices between apparently similar bonds can persist, defying arbitrage activities. Notions of segmented markets or preferred habitats are discussed, as well as empirical evidence on liquidity premia in the fixed-income markets. Section 14.3 starts with the older literature on tests of the expectations hypothesis, describes the new models distinguishing subjective expectations extracted from surveys and expectations based on statistical models, and concludes with adaptive learning, in particular the ability of learning models to fit professional forecasts, including in the context of structural models. Section 14.4 reviews term structure models recently introduced to face an ever-growing sample where the target rate is close to its lower bound in the US, Japan as well as in and around the Euro zone.

14.2 Liquidity

14.2.1 BILLS, NOTES AND BONDS

Nowadays, investors can trade more than three hundred US Treasury securities with maturities between only a few days and up to thirty years (excluding inflation-indexed securities). Their prices tend to move together, since the underlying rates used to discount each security's cash flows are highly correlated. Figure 14.1 reports estimates of the yield curve for par coupon bonds at the end of April, May, and June 2013, respectively. The yield curve steepened significantly over this period, coinciding with intense speculation as to the pace at which the FOMC would "taper" its asset purchases.

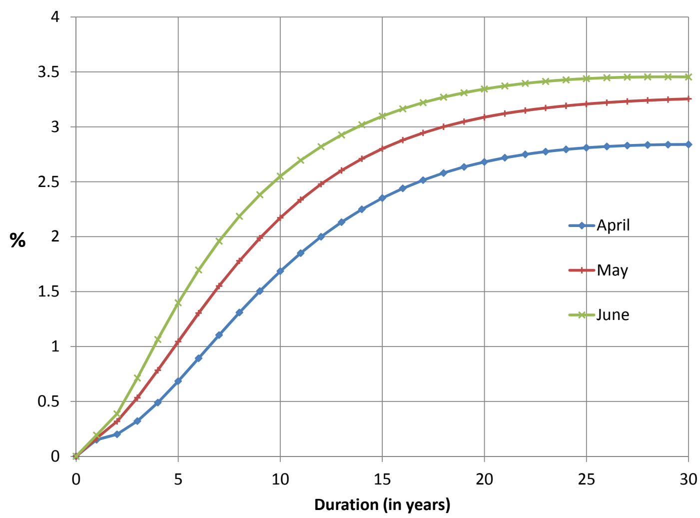


FIGURE 14.1 Yield-to-Maturity Curves.
Yields to maturity for par Treasury securities at the end of April, May and June of 2013 from the Gurkaynak, Sack and Wright (2006) data set and available from the Federal Reserve Board website.

Most importantly, Figure 14.1 illustrates the tight correlations between yields of different maturities. In fact, the entire curve of yields necessary to discount any US Treasury security can be closely approximated with a small number of risk factors—generally three—and the fit of term structure models is assessed with respect to smoothed yield curves. Nonetheless, some significant information is lost as we move between the observed yields to maturity (ytm) and the unobserved smoothed yields depicted in Figure 14.1. To illustrate, Figure 14.2 reports the ytm against duration on the x-axis, for every security in the CRSP data set and for one day at the end of March 2013.[1]

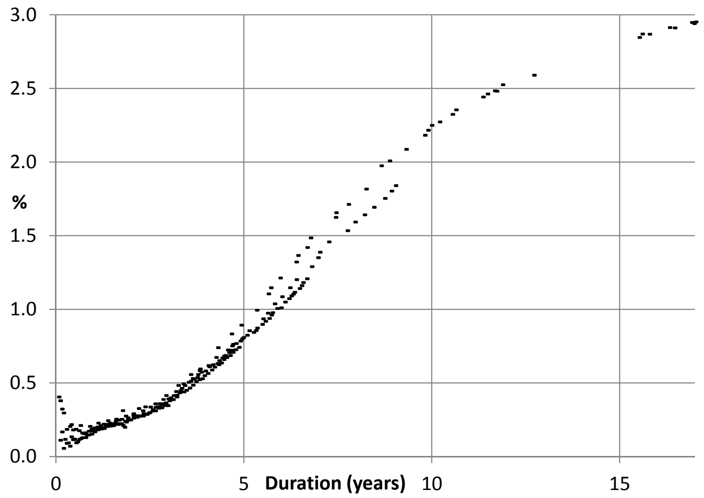


FIGURE 14.2 **US Treasury Yield-to-Maturity Curve on 28 March 2013**

Figure 14.2 a bird's eye view of what is depicted in Figure 14.1: bonds of a similar duration have a similar ytm. One could draw a curve connecting maturities, with yields ranging from 0% to 3%. But Figure 14.2 also reveals large departures from a smooth curve. There is a clear separation—a difference of around 20 basis points (bps)—between the (lower) yields of recently-issued 10-year notes and the (higher) yields of bonds originally issued with 30 years to maturity—located between years 5 and 10 on the duration axis. The separation cannot be attributed to a transitory on-the-run effect due to the recent issue being "special" in the repo (funding) market (Duffie, 1996): the difference between the recent 10-year issue and the just-off-the-run issue is only 4 bps on that day. This is also not due to differences in market liquidity: the bid-ask spreads are similar across all recent and old issues.

We observe a similar separation at the short end of the maturity spectrum between the yields on bills, which were originally issued with less than one

[1] We plot ytm against duration since, for our purpose, we want to emphasize yield differences that are not accounted for by differences in cash flows. Controlling for duration provides a first-order adjustment for coupon differences. In addition, the volatility of yields implicit in bond options was close to a multi-year low at the end of March 2013. This minimizes the effect due to differences in bond convexity.

year to maturity, and the yields on older notes and bonds. Again, the separation arises in the age dimension. The largest deviations correspond to old 30-year bonds maturing soon. For a small intermediate range of maturities—around two years to maturity—the ytm dispersion seems more compressed, probably within twice the bid-ask spreads for off-the-run issues, and appears consistent with a smooth yield curve. We can also find episodes where the yield curve ap-

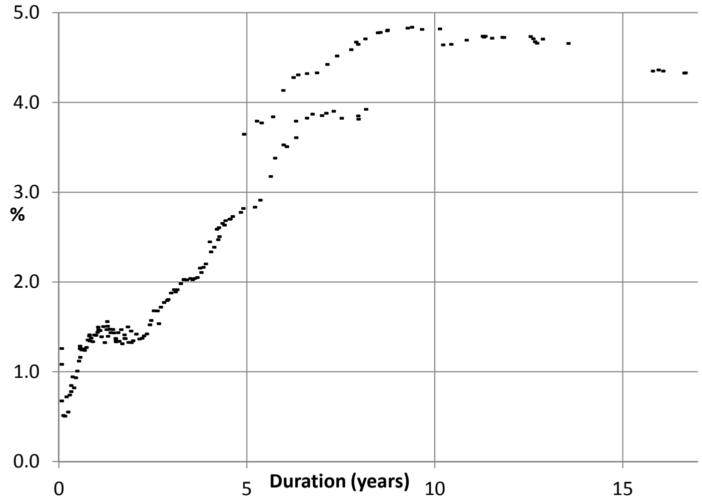


FIGURE 14.3 **US Treasury Yield-to-Maturity Curve on 31 October 2008**

pears dislocated. Figure 14.3 shows the observed ytm at the end of the October 2008, revealing what were perhaps the worst conditions in financial markets. Again, we observe a clear separation between old bonds and recent notes, but the yield difference between nearly identical securities has more than doubled, to around 50 bps. We also observe a similar separation for short maturities where, again, the deviations are much larger than in Figure 14.2.

Figure 14.3 reveals an additional pattern. There are significant troughs at regular intervals along the duration axis: the yield curve appears flat or even downward sloping at some points. These humps are typically close to maturity points with regular issuances. The most recent issue typically lies near the lowest point, but the yields of older issues located nearby are also pulled downward by arbitrage forces.

14.2.2 MARKET LIQUIDITY AND SHORT-SELLING COSTS

The market for bills appears partially segmented from the markets for old notes and bonds (Garbade, 1984; Kamara, 1994; Duffee, 1996). Similarly, the market for old bonds appears partially segmented from the market for more recently issued bonds Cornell and Shapiro (1989). Figures 14.2 and 14.3 suggest that this segmentation offers arbitrage strategies. Traders should buy the older bonds, and short the recent issues. For instance, Warga (1992) compares excess returns from holding recent issues or from holding older bonds while matching durations. In practice, the older bonds are less liquid and more costly to trade. Nonetheless, this effect is larger than bid-ask spreads and larger than the dispersion of prices quoted by different dealers for a given bond Garbade and Silber (1976). Amihud and Mendelson (1991) also find that the difference in market liquidity is not sufficient to explain the note–bills yield differential.

Amihud and Mendelson (1991) conjecture that sufficiently high costs for borrowing a security in order to sell it short could rationalize the observed yield differential. In practice, short-sellers can obtain the security from a bond holder, via a reverse repo transaction, or from the holder's custodian, via a securities lending agreement. In the case of the repo transaction, the cost of borrowing the security is given by the spread S between the "special" repo rate offered by short-sellers to borrow specific securities and the repo rate for general collateral. Duffie (1996) examines this link formally in an equilibrium where the general collateral rate, the special repo rates (or borrowing fee) and the prices of two bonds are determined jointly.[2] In equilibrium, the borrowing costs s are such that the demand for short-selling the bond $D(s)$, equals the supply by lenders of securities. In his notation,

$$D(s) = \frac{q_L S}{\bar{\lambda}}, \tag{14.1}$$

where q_L is the quantity of bonds held by potential lenders and $\frac{S}{\bar{\lambda}}$ is the fraction of lenders with opportunity costs $\lambda \leq S$. Duffie shows that the spread increases with the demand for short-selling this bond and the opportunity costs of lenders. He also shows that specialness increases with the market liquidity of the bond. Krishnamurthy (2002) provides an extension where bond holders' liquidity preference identifies a unique equilibrium.

Empirically, Jordan and Jordan (1997) for Treasury notes, and Krishnamurthy (2002) for 30-year bonds, show that borrowing costs contribute to a large extent to observed price differences between recent issues. Expanding on these results, Fontaine and Garcia (2012) extract a funding liquidity factor FL from the observed mispricings between pairs of Treasury securities across a range of 11 maturities (see also the discussion in Section 14.2.6). Expanding the range of maturities helps filter the persistent and systematic price dispersions. Following Duffie (1996), this funding factor should reflect expectations of future special repo rates across a range of maturities. To check this, Figure 14.4 reports the funding factor FL against the first principal component of *future* repo spreads to borrow on-the-run bonds with 2, 5, 10 and 30 years to maturity.

[2] The active literature on Treasury auctions is related to the discussion in this section. Several papers emphasize predictable price variations and apparent overpricing for Treasury securities following the issuance auction cycle (Jegadeesh, 1993; Cheria, Jacquier and Jarrow, 2004; Lou, D., H. Yan, and J. Zhang, 2013).

For each maturity, we average the spreads over the next three months. We then extract the first principal component PC across maturities. The relationship is clearly positive (a regression of PC on FL produces an R^2 of 28%). The relationship is statistically and economically significant. A one-standard deviation change in FL yields one-half of a standard deviation increase common across future repo spreads.

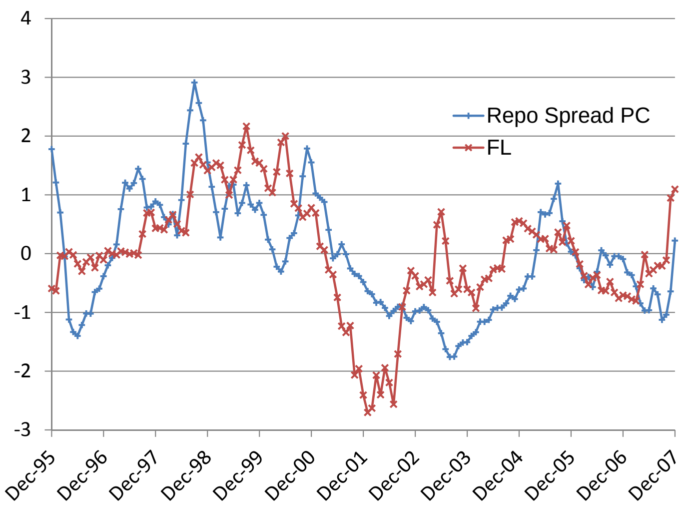


FIGURE 14.4 Funding Risk and Future Repo Spreads
Funding liquidity risk factor FL at the end of each month against the first principal component extracted from the special repo spread for bonds with maturities of 2, 5, 10, and 30 years where we use the average spread over the next three month for each maturity. Both series are normalized with a mean zero and unit standard deviation.

Duffee (1996) and Krisnamurthy (2002) take the relative liquidity of different bonds as given. But why are older bonds less liquid? Garbade and Silber (1976) argue that "seasoning [...] measures whether the issue is trading actively because it has not yet been 'locked away' in investor portfolios." See also Sarig and Warga (1989); Amihud and Mendelson (1991); and Vayanos and Weill (2006). Warga (1992) confirms that excess returns of the arbitrage strategy increase when the age of the seasoned bonds increases. Can this clientele effect rationalize at once the observations that recent issues have higher market liquidity, higher prices and higher borrowing costs?

Vayanos and Weill (2006) provide a positive answer. Liquidity and specialness are not independent explanations. Since a short-seller must eventually buy back the security borrowed, the short-sellers concentrate activity around the security that is easier to locate, which is the same that other short-sellers are borrowing. Today's short-sellers are tomorrow's buyers, increasing the liquidity and the price of the bond. The combination of short-selling and search frictions generates asymmetries between the equilibrium price *and* liquidity of each asset. Adding search in the repo market, Vayanos and Weill (2006) also generate lower special repo rates (higher borrowing costs), adding to the price difference. As in most equilibrium with search frictions, the value of a trade is shared between a buyer and a seller according to an exogenous "bargaining power" parameter. Empirically, Banerjee and Graveline (1996) decompose the premium between on-the-run and just-off-the-run 5-year and 10-year notes and estimate that, from November 1995 through July 2009, short-sellers bore 22% and 37% of the premium, respectively. Consistent with existing results, they also find that profits from the arbitrage strategies are quite volatile but that they are statistically different than zero.

14.2.3 HEDGING DEMAND

The demand for hedging duration risk is perhaps the most important determinant of the short-selling demand in the Treasury market. Indeed, Graveline and McBrady (2011) find that proxies for hedging demand and proxies for buy-and-hold investors' demand to own liquid securities explain roughly the same share of repo spread variations. Much of the duration risk variation comes from the interest rate convexity of mortgage-backed securities (MBS).[3] Fernald, Keane and Mosser (1994) detail how higher current or expected interest rates reduce the mortgage prepayment and refinancing rates, increase the duration of MBS portfolios, and increase the hedging demand of dealers and investors with large MBS inventories. As an illustration, Figure 14.5 compares Barclay's index of MBS duration with the notional amount of derivatives held by Fannie Mae for risk management purposes between 2000 and 2012.[4] The two series are clearly correlated and the changes in Fannie Mae's hedging demand are economically large, varying between 400 and 1,200 billion US dollars.

As made clear in Vayanos and Weill (2006), the increase in short-selling demand makes the limits-to-arbitrage more severe in the Treasury market, increasing the spread between new and old bonds. In addition, theoretical results link the price of interest rate risk to changes in aggregate duration (Hanson, 2014; Malkhozov, Mueller, Vedolin and Venter, 2014). Again, the spread between new and old bonds then takes a new significance: it contains information about the compensation for interest rate risk. In addition, the same authors show that duration risk predicts bond returns and this information is not subsumed by factors from the (smoothed) yield curve. [5]

[3] Hedging interest rate futures or options risk is another source of short interest in the Treasury market.

[4] We thank Aytek Malkhozov for sharing these data.

[5] Gabaix, Krishnamurthy and Vigneron (2007) also show that prepayment risk carries a positive risk premium in the MBS market itself. These results add to the documented effect of MBS duration on option-implied and realized volatility of bond yields (Perli and Sack, 2003; Duarte, 2008)

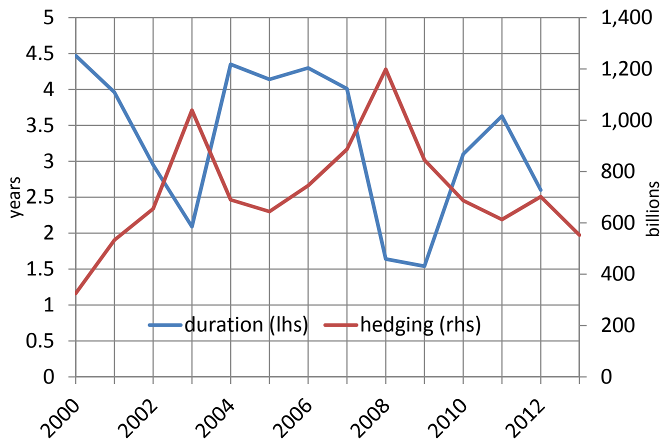


FIGURE 14.5 **MBS Duration and Fannie Mae's Hedging**
Barclay's MBS duration index (in years) and the Fannie Mae notional amount of derivatives used for risk management purposes (billions USD). Annual data, 2000–12.

14.2.4 RISKY ARBITRAGE

Banerjee and Graveline (2013) find that "variation in the repo premium does not completely account for variation in the price premium" (p.650). In addition, Figure 14.2 shows that off-the-run notes initially issued with 10 years to maturity have a lower yield than old bonds initially issued with 30 years to maturity. These two securities have different market liquidity, explaining part of the yield difference, but they carry the same borrowing costs on the repo market (excluding the few most recent issues). Similarly, Longstaff (2004) shows that the spread between the yield on RefCorp bonds and seasoned Treasury bonds with identical cash flows is not commensurate with the lower bid-ask spreads and lower repo rate of Treasury issues.

The evidence is inconsistent with the prediction that the price differences between similar bonds should compensate for the expected transaction costs and borrowing costs until bond prices converge (see, e.g., Proposition 1 in Duffie, 1996 or Equation (8) in Krishnamurty 2002). But these models do not consider the effect of risk. First, rolling over overnight short positions is risky. For instance, Bartolini et al. (2011) find that the general collateral repo rates of Treasury securities vary significantly, while the (higher) repo rates on agency securities and MBS collateral are not meaningfully different from the unsecured rates. Buraschi and Menini (2002) ask whether *term* special repo rates correctly predict future *overnight* special repo spreads. This is an analog to tests of the expectations hypothesis between long-term rates and the short-term interest rate.

They conclude that forward spreads implicit in term repo spreads overestimate changes in future overnight repo spreads. The effect is significant and largely explained by the volatility of the overnight special repo rate. In other words, rolling over a short position generates an exposure to an unexpectedly large aggregate short interest rate in the future.

Figure 14.6 illustrates the large variability of the repo spread in transactions involving the on-the-run 10-year Note. The repo spread can be large, often as large as the underlying general collateral repo rate. More importantly, its volatility can be low at times, but with rapid and persistent increases. The bottom panel of Figure 14.6 displays the conditional volatility from an EGARCH(1,1) model fitted to the data. Estimates confirm a strong leverage effect: periods of high volatility are also periods of high repo spreads, increasing the risk associated with short-selling securities. Liu and Longstaff (2004) find that the optimal strategy between new bonds and old bonds rarely fully exploits the arbitrage, since the speculator's final wealth is inversely related to the probability of hitting the speculator's borrowing constraint.

14.2.5 SEGMENTED MARKETS AND PREFERRED HABITATS

The clientele demand for new and old bonds is similar in spirit to the view that investors have "preferred habitats" (Culbertson, 1956; Modigliani and Sutch 1967). The clientele demand may be scattered across bond maturities, but it can also be scattered across the illiquidity spectrum—as in Krishnamurthy (2002)—across bonds with different ages. The effect of preferred habitats is analyzed formally in Vayanos and Vila (2009), where investors have clienteles' preference for specific maturities. Empirically, Chen et al. (2014) find that the aggregate portfolios of insurance firms exhibit a liability habitat—driven by the need to immunize the interest rate risk of operating liabilities—and a horizon habitat—due to the preference for holding risk-free securities over a specific investment horizon.

In Vayanos and Vila (2009), risk-averse arbitrageurs integrate the maturity-specific markets, eliminating risk-free arbitrage. However, the risk premium required by arbitrageurs propagates local clientele shocks across the term structure of yields. Building on this prediction, Greenwood and Vayanos (2014) show that a maturity-weighted measure of debt-to-GDP is positively related to bond returns and argue that the effect is stronger when intermediaries' wealth has declined. More generally, interest rate risk increases when intermediaries wealth' decreases (or if their risk aversion increases). Guibaud, Nosbusch and Vayanos (2013) study the optimal maturity structure of government debt within a clientele-based equilibrium model of the yield curve. Consistent with their model, they find that the median age of the population is correlated with the slope between 10- and 30-year bonds in the cross-section of OECD countries. Countries with older populations have relatively more investors with shorter investment horizons.

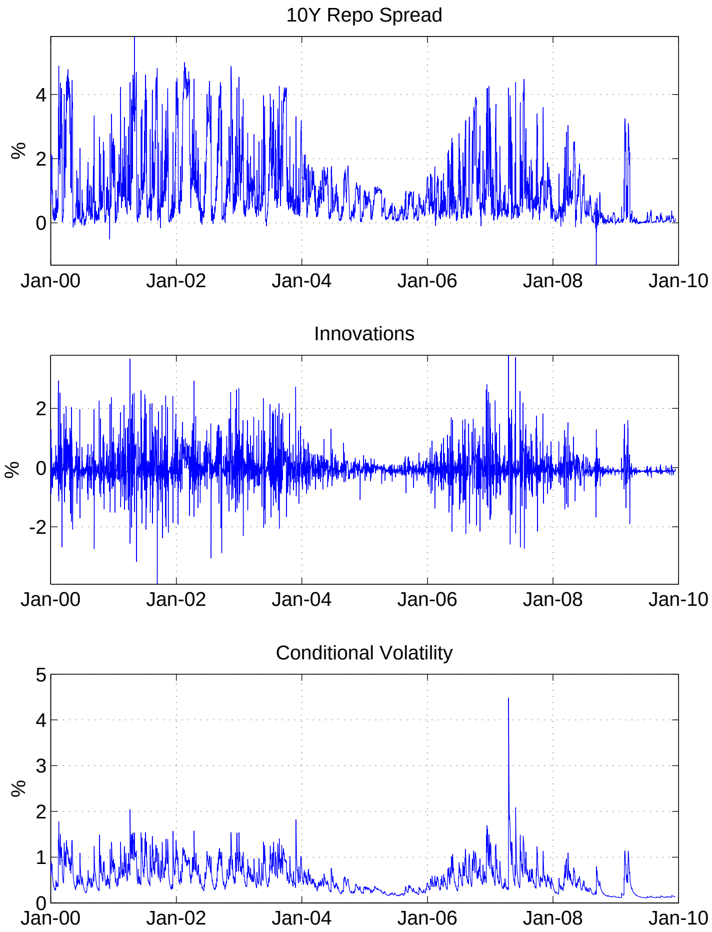


FIGURE 14.6 **10-Year Repo Spreads**

Repo spreads for transactions involving the on-the-run 10-year note (top panel). Innovations from fitting an ARMA(1,1) model to the repo spreads (middle panel). Estimated conditional volatility from fitting an EGARCH(1,1) model to the innovations (bottom panel). All series in percentage points. Daily data from ICAP, November 1995 to July 2010.

14.2.6 FUNDING RISK

Funding risk provides a channel between conditions in the repo markets, limits-to-arbitrage and the bond risk premiums. Krishnamurthy (2002) shows that the expected repo spreads and bond spreads increase when the clientele's preference for the more-liquid assets increases. Adrian and Shin (2008) show that broker-dealers' leverage is procyclical and that the repo market is the marginal source of broker-dealers' funds. The reduction of leverage also implies that dealers are less willing to lend bonds or funds for the purpose of the arbitrage strategy. The evidence suggests that shocks to intermediaries' or to bondholders' liquidity preferences lead to wider *apparent* arbitrage opportunities across Treasury securities. Summing up, the spread between new and old bonds then takes on an important significance: it contains information about intermediaries' funding risk.[6]

Fontaine and Garcia (2012) provide direct evidence that measures of limits-to-arbitrage from the Treasury market can predict bond excess returns. They estimate a standard affine term structure model using observed coupon bond prices instead of the smoothed zero-coupon yields. They add a funding risk factor discounting the price of older bonds,

$$P(F_t, L_t, Z_{n,t}) = \sum_{m=1}^{M_n} D_t(m) \times C_{n,t}(m) + \zeta(L_t, Z_{n,t}), \tag{14.2}$$

where $Z_{n,t}$ summarizes any security-specific characteristics. The specification of $\zeta(L_t, Z_{n,t})$ in Fontaine and Garcia (2012) relies on the observation that older bonds become less liquid and less expensive. The price premium is given by

$$\zeta(L_t, Z_{n,t}) = L_t \times \beta_{M_n} \exp\left(-\frac{1}{\kappa}\, age_{n,t}\right), \tag{14.3}$$

where $age_{n,t}$ is the age, in years, of the bond at time t and β_M controls the average premium at each fixed maturity M. The filtered estimates of funding risk capture the price differences observed in Figure 14.2 that are (i) persistent, (ii) common across maturities, and (iii) correlated with age. As predicted by theory, the funding risk factor is correlated with measures of liquidity: it is correlated with current bid-ask spreads and it predicts future bid-ask spreads. Funding risk predicts future repo spreads. However, the funding risk factor is also correlated with broader liquidity indicators, including the shape of the yield curve; investment flows into money-market mutual funds; banks' non-borrowed reserves at the Fed; the growth of the shadow banking sector; and changes in monetary aggregates. Consistent with Krishnamurthy (2002), the dispersion of prices in Figures 14.2 and 14.3 is associated with investors' liquidity preference.

[6]This is also analogous to results in Vayanos and Vila (2009) where the limited arbitrageur's ability to integrate segmented markets generates a risk premium.

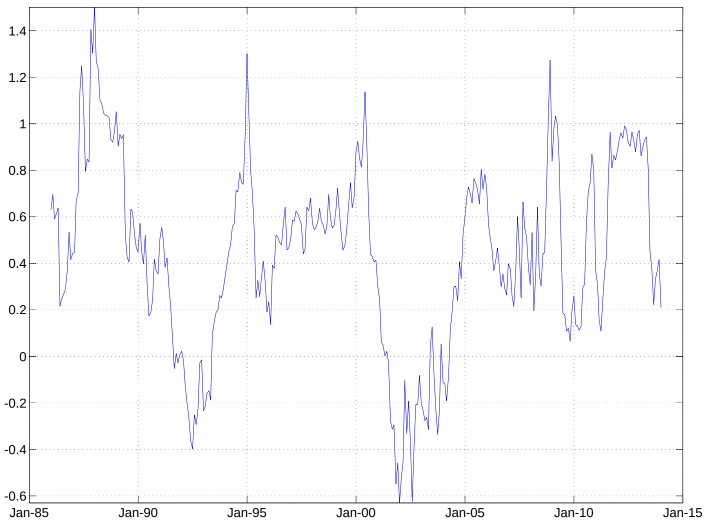


FIGURE 14.7 Funding Liquidity Risk

Filtered estimates of the funding risk factor L, as in Fontaine and Garcia (2012), but normalized to a mean of zero and a variance of 1, and updated in a monthly sample from January 1986 to December 2013.

Fontaine and Garcia (2012) then check whether the funding risk factor predicts the risk premium across fixed-income markets. Figure 14.7 reproduce their factor, but updated to a more recent sample. Conditions in the US funding market exhibit a spike during the October 1987 crash, the Mexican peso crisis, toward the end of the millennium, and in 2005 when Ford and GM were downgraded to junk status. More recently, conditions spiked in 2008, in the summer of 2010, and throughout most of the sovereign crisis in Europe.

Table 14.1 reproduces some of the predictability results in Fontaine and Garcia (2012) (see their Table 8B), and reports the R^2 and coefficients on the funding risk factor in the following regressions of annual Treasury bond returns on funding risk:

$$xr_{t+12}^{(m)} = \alpha^{(m)} + \delta^{(m)} L_t + \beta^{(m)T} f_t + \epsilon_{(t+12)}^{(m)}, \tag{14.4}$$

where $xr_{t+h}^{(m)}$ are the annual excess returns on a zero-coupon bond with m years to maturity (from CRSP data), L_t is the liquidity factor and f_t is a vector of annual forward rates $f_t^{(h)}$ from one to five years. We include f_t to control for the information content of forward rates.

The expected return on Treasury bonds is lower when the funding risk factor increases. This is consistent with an increase of lenders' or bondholders' liquidity preference driving up prices. The effect is economically significant: a one-standard deviation increase of the funding risk factor lowers expected returns by 0.8% and 2.7% for a bond with two and five years to maturity, respectively. The R^2s are doubled compared with the ones obtained when excluding the funding risk factor (unreported). The results cannot be attributed to the 2008–09 crisis: the regressions are estimated using data between 1986 and 2007. Adding data from the crisis period only strengthens the results.

Table 14.1 Treasury Excess Returns and Funding Risk

Maturity	α	$f_t^{(1)}$	$f_t^{(2)}$	$f_t^{(3)}$	$f_t^{(4)}$	$f_t^{(5)}$	L_t	R^2
2y	0.72	0.29	-1.31	1.88	0.93	-0.95	-0.78	41.65
	(3.49)	(0.49)	(-1.18)	(1.50)	(1.04)	(-1.60)	(-5.97)	
3y	1.31	0.15	-2.26	4.32	0.76	-1.49	-1.55	41.66
	(3.41)	(0.14)	(-1.13)	(1.89)	(0.48)	(-1.27)	(-5.93)	
4y	1.79	-0.51	-1.74	4.58	1.53	-1.85	-2.18	42.82
	(3.53)	(-0.35)	(-0.66)	(1.51)	(0.75)	(-1.13)	(-6.07)	
5y	1.98	-1.51	-0.24	4.57	0.36	-0.81	-2.66	40.87
	(3.23)	(-0.84)	(-0.07)	(1.24)	(0.15)	(-0.39)	(-5.83)	

Results from the regressions of annual excess returns $xr_{t+h}^{(m)}$ on a zero-coupon bond with m years to maturity,

$$xr_{t+12}^{(m)} = \alpha^{(m)} + \delta^{(m)} L_t + \beta^{(m)T} f_t + \epsilon_{(t+12)}^{(m)},$$

where f_t is a vector of annual forward rates $f_t^{(h)}$ from one to five years, and L_t is the liquidity factor. Regressors are demeaned and divided by their standard deviations. Newey-West t-statistics (in parentheses) with 15 lags. End-of-month CRSP data (1985:12–2007:12).

Fontaine and Garcia (2012) also find that an increase in the funding risk factor predicts higher returns across the LIBOR, swap, agency and corporate bond markets. Importantly, the estimated predictability coefficients follow a flight-to-liquidity pattern. Limits-to-arbitrage in the Treasury market reflect liquidity preferences and funding risk, transmitting shocks to the risk premium across markets. Minimally, the results show that shocks to intermediaries' willingness or ability to facilitate arbitrage activities are correlated to shocks driving the risk premium across markets. In addition, these shocks potentially cause some of the risk premium variations.

Because repo markets play an essential role in the provision of liquidity across the fixed-income market, the evidence strongly suggests that intermediaries amplify and propagate those shocks when managing their funding risk. Fontaine, Garcia and Gungor (2015) show that funding risk is transmitted to other markets. They find that funding shocks generate cross-sectional variations in liquidity and volatility; that funding shocks are correlated with contemporaneous returns in the cross-section of stock returns; and that this risk is priced across a broad set of test portfolios.

14.2.7 IMPLICATION FOR TERM STRUCTURE MODELS

The predictive power of the funding risk factor is robust to the inclusion of forward rates. In fact, the information content of these two factors appears largely independent. In other words, observed bond prices contain information about the future risk premium that is not spanned in the smoothed zero-coupon yield curve.[7]

Funding shocks are correlated with contemporaneous bond returns. We estimate a VAR(1) for the three term structure factors and the funding factor, jointly, and we compute the residuals based on the estimated dynamics. Table 14.2 reports the results from contemporaneous regressions of monthly bond returns on funding and term structure shocks:

$$r_{t+1}^{(m)} = \alpha + \beta_u^\top u_{t+1} + \beta_X^\top X_t + \epsilon_{t+1}, \tag{14.5}$$

where $r_{t+1}^{(m)}$ is the monthly log-returns from an off-the-run zero-coupon bond with maturity m, $X_t = [L_t \; F_t]$ stacks the funding liquidity and term structure factors from the no-arbitrage model with liquidity, and u_{t+1} is the Cholesky decomposition of $X_{t+1} - E_t[X_{t+1}]$ computed from a VAR(1). We include lagged funding and term structure factors as conditioning information to increase precision.

The evidence is unambiguous. Funding shocks are associated with positive bond returns contemporaneously. For a 10-year bond, a funding shock is associated with 0.3 percent monthly returns. Consistent with predictability regressions, this higher current value of funding liquidity is associated with lower expected returns in the future. This holds for every maturity. Coefficients of lagged term structure factors and of their innovations have the expected sign.

We identify the shocks via a Cholesky decomposition of the estimated covariance matrix where funding risk is ordered first. A structural interpretation suggests that an economically significant share of the term structure factor innovations is attributable to fundin,g shocks. For instance, this will be the case if Treasury bonds provide a hedge against funding liquidity shocks. Overall, the evidence points to the important role of intermediation frictions in the level and variability of yields.

14.3 Learning

The expectations hypothesis is certainly a good point of departure to study the term structure of yields and the predictability of bond returns. Indeed, testing

[7] The principal components (PCs) computed from the Fama-Bliss CRSP files (T-bills rates at maturities of 1, 2, 3, 4, 5 and 6 months; zero-coupon yields at maturities of 1, 2, 3, 4 and 5 years) span only 47% of funding liquidity variations. PCs from the data set of Gurkaynak, Sack and Wright (2006), computed using maturities of 3, 6, 9 and 12 months as well as 1, 2, 3, 4, 5, 7, 10 years, span 54% of the funding liquidity variations. See Joslin, Priebsch, and Singleton (2014) for a discussion of unspanned risk within dynamic no-arbitrage term structure models.

Table 14.2 **Funding Liquidity and Contemporaneous Treasury Bond Returns**

	3m	**6m**	**9m**	**1y**	**2y**	**3y**	**4y**	**5y**	**7y**	**10y**
α	0.41	0.43	0.45	0.46	0.52	0.57	0.61	0.65	0.71	0.78
	Funding Liquidity Innovations									
$u_{liq,t+1}$	0.011	0.013	0.008	0.004	0.017	0.057	0.103	0.145	0.216	0.299
	(5.746)	(7.319)	(3.932)	(1.749)	(4.647)	(10.530)	(11.317)	(12.525)	(15.569)	(6.688)
	Nelson-Siegel Factor Innovations									
$u_{lvl,t+1}$	-0.006	-0.031	-0.070	-0.119	-0.370	-0.658	-0.953	-1.246	-1.806	-2.571
$u_{slp,t+1}$	-0.028	-0.076	-0.123	-0.164	-0.278	-0.337	-0.368	-0.384	-0.395	-0.395
$u_{crv,t+1}$	-0.003	-0.031	-0.073	-0.123	-0.336	-0.520	-0.653	-0.736	-0.785	-0.680
	Funding Liquidity									
L_t	-0.018	-0.025	-0.038	-0.054	-0.131	-0.204	-0.266	-0.316	-0.392	-0.467
	(-10.091)	(-14.852)	(-21.013)	(-25.969)	(-45.261)	(-44.930)	(-37.740)	(-34.771)	(-30.051)	(-13.147)
	Nelson-Siegel Factors									
$F_{1,t}$	0.143	0.160	0.175	0.190	0.250	0.306	0.355	0.398	0.476	0.584
$F_{2,t}$	0.146	0.148	0.147	0.146	0.143	0.137	0.127	0.116	0.094	0.074
$F_{3,t}$	0.003	0.004	0.006	0.010	0.033	0.060	0.087	0.111	0.149	0.184
R^2	97.7	98.3	98.4	98.6	99.3	99.4	99.4	99.4	99.2	97.6

Results from the contemporaneous regressions,

$$r_{t+1}^{(m)} = \alpha + \beta_u^\top u_{t+1} + \beta_X^\top X_t + \epsilon_{t+1}$$

$r_{t+1}^{(m)}$ is the monthly log-returns from an off-the-run zero-coupon bond with maturity m, $X_t = [L_t\ F_t]$ stacks the funding liquidity and term structure factors from the AFENS model with liquidity, and where u_{t+1} is the Cholesky decomposition of $X_{t+1} - E_t[X_{t+1}]$. Regressors are demeaned and divided by their standard deviations. Newey-West standard errors (3 lags) in parentheses. End-of-month CRSP data (1985:12–2007:12).

the EH was a major concern for researchers in the eighties. To set the problem clearly, one can start with a fundamental asset pricing equation for bond excess returns:

$$E_t^*[xR_{t+1}^{(n)}M_{t+1}] = 0, \tag{14.6}$$

where $xR_{t+1}^{(n)}$ is the returns from an n-period bond between t and $t+1$, M_{t+1} denotes the marginal utility of an investor, and the expectation $E_t^*[\cdot]$ is taken under the subjective beliefs of the investor (denoted by *). Applying the covariance formula and introducing a expected excess return under the true probability measure, $E_t[xR_{t+1}^{(n)}]$, one obtains

$$E_t[xR_{t+1}^{(n)}] = \Big(E_t[xR_{t+1}^{(n)}] - E_t^*[xR_{t+1}^{(n)}]\Big) - \frac{\text{cov}_t^*[M_{t+1}, xR_{t+1}^{(n)}]}{E_t^* M_{t+1}}. \tag{14.7}$$

The first term on the right-hand side captures the difference between the expected excess returns computed using the investor's beliefs or using the true measure. The second term captures the risk premium. The larger part of the literature assumes that the first term is zero, since investors are typically endowed with full-information rational expectations. The expectations hypothesis implies that the risk premium is either zero or constant through time. Therefore, predictability regressions of bond returns—the workhorse in this literature—provide only a test of the joint hypothesis that investors' expectations conformed with the expectations hypothesis and with full-information rational expectations. The literature provides a quite different interpretations of the evidence that bond returns were predictable. Most authors argue that time-varying term premia are responsible for the predictability, but others argue for the importance of expectation errors, invoking the underreaction of long rates (or expected future short rates).[8]

14.3.1 YIELD SURVEY FORECASTS

Some authors avoid the conundrum by relying on survey data on interest rate expectations. From the EH, investors' forecasts should be unbiased, efficient, consistent, and forecast errors should be orthogonal with respect to freely available information. Strikingly, the use of survey data as a means to test the EH makes an appearance as early as Meiselman (1962). Friedman (1980) uses the survey conducted in the *Goldsmith and Nagan Bond and Money Market Letter* among a selected panel of market professionals. The data provide the median response for one-quarter-ahead and two-quarter-ahead predictions of several interest rates. Friedman (1980) concludes that survey forecasts were biased, that they failed to be consistent and that they did not efficiently exploit the information in past interest rate movements. Moreover, he finds that common current macroeconomic and macro-policy variables forecast ex-post prediction errors for long-term rates (but not for short rates).

[8] See Shiller and McCulloch (1987) for a survey of this literature.

Froot (1989) also uses the survey conducted in the *Goldsmith and Nagan Bond and Money Market Letter*. He runs two sets of tests with survey data. First, he constructs a direct test of the EH by regressing survey forecasts of yields on short- and long-term instruments on forward premia:

$$, s_{t+j}^{(k-j)} = \alpha + \beta f p_t^{(j,k)} + \epsilon_{t,j} \tag{14.8}$$

where

- $s_{t+j}^{(k-j)}$ is the survey expected yield on a $k-j$-period bill, purchased at time $(t+j)$ less the yield at time t on a j-period bill, purchased at time t,
- and $fp_t^{(j,k)} = f_t^{(j,k)} - i_t^{(j)}$ is the forward premium at time t on a $k-j$-period bill, j periods into the future, less the yield on a j-period bill.

The forecast $s_{t+j}^{(k-j)}$ is a median survey measure and it measures the spread with error. The test verifies whether the biased prediction is attributable to a time-varying term premium. Froot (1989) concludes that the expectations theory on short-term instruments is rejected, since the regression coefficient β is statistically less than one. However, the survey data tend to show that expected long-rate changes conform to the expectations theory: changes in the spread are reflected one-for-one in changes in expected future long rates (β is statistically not different from one). So it appears that term premia are more important in short-duration instruments.

A second test involves expectational errors—that is, the survey prediction error of the spread $i_{t+j}^{(k-j)} - i_t^{(j)}$. Froot (1989) runs regressions of these errors on the forward premium defined in (14.8). For short maturities, expectational errors appear to be unsystematic. For longer-maturity instruments, results show that expectational errors are systematic, which explains the poor predictive power of the spread for future long-rate changes.

Recently, Cieslak and Povala (2013) have shown that the expected real federal funds rate implied by surveys differs from its statistical counterpart obtained under the assumption of full-information rational expectations. Figure 14.8 compares forecasts of the federal funds rate based on survey data—labeled $r_t^{e,surv}$—and based on a projection of realized $t+1$ federal funds rates on lagged instruments—labeled $\hat{r}_t^{e,FIRE}$.[9] The results show that survey forecasts miss both downturns and upturns. In particular, forecasts of future rates are systematically above statistical forecasts throughout all recessions but one. Figure 14.9 reports the difference between the forecasts $r_t^{e,FIRE}$ and $\hat{r}_t^{e,surv}$, which Cieslak and Povala (2013) label $MP^{\perp}$. They show that $MP^{\perp}$ is persistent at the business cycle frequency, it becomes notably large at the beginning of recessions, and it is closely correlated with errors in unemployment forecast errors from surveys. Important to our discussion, they find that several

[9] $FIRE$ stands for full-information rational expectations. As instruments, Cieslak and Povala (2013) use time-t year-on-year CPI inflation, federal funds rate, annual change in the rate of unemploymfent, one-year nominal yield, and the term spread. They also include the $t-1$ term spread. We thank Anna Cieslak for sharing these results.

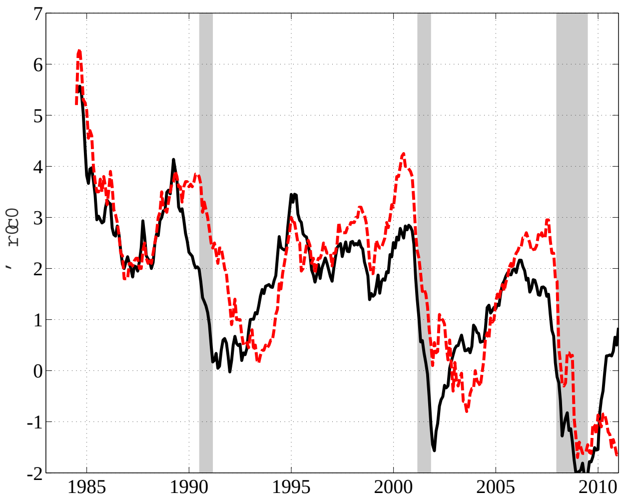


FIGURE 14.8 **Ex-ante real fed funds rate from Cieslak and Povala (2013)**

conditioning variables used in predictive regressions of bond returns, especially variables related to the real activity, predict the unexpected returns associated with $MP^{\perp}$ and that these are essentially uncorrelated with the survey-implied risk premia.

Recent work provides additional evidence against the rationality of survey forecasts. Gourinchas and Tornell (2004) use forecasts of 3-month euro rates, 3, 6 and 12 months ahead, for the G-7 countries.[10] Jongen, Verschoor and Wolff (2005) use survey data on 3-month maturity euro, interbank, and T-bill rates.[11] As in Froot (1989), they find that survey forecast errors are linked to the forward premium. Bacchetta, Mertens and van Wincoop (2009) find systematic evidence that survey forecasts are predictable across markets, across sample periods and across countries.[12] Coibion and Gorodnichenko (2012) document that mean forecasts in surveys of consumers, firms, central bankers and professional

[10]The data are based on the Financial Times Currency Forecaster, with 48 respondents from both large banks and multinational companies active in the foreign exchange market.

[11]Jongen, Verschoor and Wolff (2005) use data from the Consensus Economics of London on 3-month-ahead expectations of interest rates in 20 industrialized countries by 250 professional forecasters worldwide.

[12]Bacchetta, Mertens and van Wincoop (2009) use the survey of exchange rate and interest rate expectations by Forecasts Unlimited Inc. For interest rates, the survey provides the expectations of 3-month LIBOR, 12-month LIBOR and 10-year government bond yields 3, 6, and 12 months ahead.

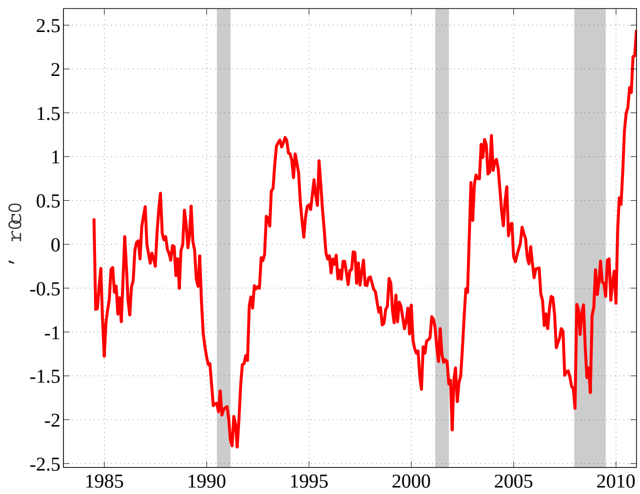


FIGURE 14.9 **Difference $MP^{\perp}$ between survey-expectations $r_t^{e,FIRE}$ and estimated rational-expectations $\hat{r}_t^{e,surv}$ forecasts of real fed funds rate from Cieslak and Povala (2013).**

forecasters do not adjust fully after shocks. These deviations are significant, both statistically and economically, and they are consistent with the predictions of models in which agents face information constraints.

14.3.2 AFFINE TERM STRUCTURE MODELS

The modern term structure literature has mostly proceeded on the assumption that predictable variations in bond returns come from risk premia. This section briefly reviews the risk premium channel in the context of benchmark dynamic term structure models (DTSMs), while subsequent sections will discuss (i) recent attempts to use the information in survey data and (ii) attempts to address the limitations of survey forecasts.

Bond Prices As usual, we start from a set of latent risk factors $\mathcal{Z}_t \in \mathbf{R}^{\mathcal{N}_Z}$. Following Feunou and Fontaine (2014), we will construct a conditional mean DTSM (CM-DTSM) based on the conditional mean of $\mathcal{Z}_t$ under the (true) historical measure $\mathbb{P}$:

$$\mathcal{E}_t \equiv E_t[\mathcal{Z}_{t+1}], \tag{14.9}$$

where we adopt a forward-looking specification of the one-period nominal interest rate i_t:

$$i_t = \rho_0 + \rho_1' \mathcal{E}_t. \tag{14.10}$$

Using a specification based on $\mathcal{E}_t$ will make the connection with survey forecasts more transparent. A standard formulation would use $\mathcal{Z}_t$ instead of $\mathcal{E}_t$ in the short rate i. This is innocuous at this point: simply neglect $\mathcal{Z}_t$, take $\mathcal{E}_t$ as the latent risk factor and follow (say) Joslin, Singleton and Zhu (2011) to derive the pricing equation.

We leave the dynamics of $\mathcal{E}_t$ (and $\mathcal{Z}_t$) under $\mathbb{P}$ unspecified for now. At this point, we need the dynamics of $\mathcal{E}_t$ under the risk-neutral measure $\mathbb{Q}$ to derive the pricing equation for yields. For tractability, we use a Markovian specification of order one:

$$\Delta\mathcal{E}_{t+1} = K_0^{\mathbb{Q}} + K_1^{\mathbb{Q}}\mathcal{E}_t + \Sigma_{\mathcal{E}}\epsilon_{t+1}^{\mathbb{Q}}, \tag{14.11}$$

where $\epsilon_{t+1}^{\mathbb{Q}}$ is a Gaussian white noise. Proposition 1 in Feunou and Fontaine (2014) translate the identification assumptions from Joslin, Singleton and Zhu (2011) to the dynamics for $\mathcal{E}_t$. The n-period nominal bond price is defined by

$$P_t^{(n)} \equiv E_t^Q\left[\exp\left\{-\sum_{j=0}^{n-1} r_{t+j}\right\}\right] = E_t^Q\left[\exp\left\{-r_t\right\}P_{t+1}^{(n-1)}\right].$$

Guess that $P_t^{(n)} = \exp\left(A_n + B_n'\mathcal{E}_t\right)$—using the fact that the model is affine—and substitute for $P_{t+1}^{(n-1)}$ to solve for the coefficients A_n and B_n:

$$\begin{aligned} B_{n+1} =& (I_{\mathcal{N}_Z} + K_1^{\mathbb{Q}\prime})B_n - \rho_1 \\ A_{n+1} =& A_n + K_0^{\mathbb{Q}\prime}B_n + \frac{1}{2}B_n'\Sigma_{\mathcal{E}}\Sigma_{\mathcal{E}}'B_n - \rho_0. \end{aligned} \tag{14.12}$$

It follows that yields are given by $y_t^{(n)} = -\frac{1}{n}\log\left(P_t^{(n)}\right) = a_n + b_n'\mathcal{E}_t$. One important implication is that yields span time-t forecasts of the future risk factors $\mathcal{Z}_{t+1}$ but that they do not span Z_t unless $\mathcal{E}_t$ and $\mathcal{Z}_t$ are linearly dependent.

Historical Dynamics The change of measure ξ_t is assumed to be exponential-affine:

$$\xi_{t+1} \equiv \frac{\exp(\lambda_t'\epsilon_{t+1}^{\mathbb{Q}})}{E_t^{\mathbb{P}}[\exp(\lambda_t'\epsilon_{t+1}^{\mathbb{Q}})]}, \tag{14.13}$$

which guarantees that $\xi_{t+1} > 0$ and $E_t^{\mathbb{P}}[\xi_{t+1}] = 1$, as required by the absence of arbitrage. This change of measure also implies that the (log-) stochastic discount factor is given by

$$m_t = \log M_t = \exp(-i_t)\xi_{t+1} = -i_t - \frac{1}{2}\lambda_t'\lambda_t + \lambda_t'\epsilon_{t+1}^{\mathbb{Q}}. \tag{14.14}$$

We derive the historical dynamics from the conditional Laplace transform of the innovations:

$$\begin{aligned} E_t^{\mathbb{P}}\left[\exp(u'\epsilon_{t+1}^{\mathbb{Q}})\right] &= E_t^{\mathbb{Q}}[\frac{\exp(u'\epsilon_{t+1}^{\mathbb{Q}})}{\xi_{t+1}}] \\ &= E_t^{\mathbb{P}}[\exp(\lambda_t'\epsilon_{t+1}^{\mathbb{Q}})]\exp\left(-\frac{1}{2}(u-\lambda_t)'(u-\lambda_t)\right) \\ &= \exp\left(-\frac{1}{2}u'u + u'\lambda_t\right), \end{aligned} \tag{14.15}$$

where the first line in 14.15 follows, since ξ_{t+1} is the change of measure between $\mathbb{Q}$ and $\mathbb{P}$, while the second line uses the definition in 14.13.[13] This conditional Laplace transform of $\epsilon_{t+1}^{\mathbb{Q}}$ under $\mathbb{P}$ defines a Gaussian distribution with unit variance but with a conditional mean given by

$$E_t^{\mathbb{P}}[\epsilon_{t+1}^{\mathbb{Q}}] = \lambda_t. \tag{14.16}$$

Assuming that the $\mathcal{N}_{\mathcal{Z}} \times 1$ vector of prices of risk λ_t is also affine,

$$\lambda_t \equiv \tilde{\lambda}_0 + \tilde{\lambda}_1 \mathcal{E}_t, \tag{14.17}$$

and using Equations 14.11 and 14.16 to compute $E_t^{\mathbb{P}}[\Delta\mathcal{E}_{t+1}]$, it follows that the dynamics of $\mathcal{E}_{t+1}$ under the historical measure $\mathbb{P}$ is given by

$$\Delta\mathcal{E}_{t+1} = K_0^{\mathbb{P}} + K_1^{\mathbb{P}}\mathcal{E}_t + \Sigma_{\mathcal{E}}\epsilon_{t+1}^{\mathbb{P}}, \tag{14.18}$$

with parameters given by

$$K_0^{\mathbb{P}} = K_0^{\mathbb{Q}} - \Sigma_{\mathcal{E}}\tilde{\lambda}_0 \tag{14.19}$$

$$K_1^{\mathbb{P}} = K_1^{\mathbb{Q}} - \Sigma_{\mathcal{E}}\tilde{\lambda}_1. \tag{14.20}$$

Dai and Singleton (2002) and Duffee (2002) show that essentially affine Gaussian term structure models are able to explain stylized facts of the bond risk premium reported by Fama and Bliss (1987) and Campbell and Shiller (1989), whereby projections of bond returns on the slope of the yield curve produce coefficients that contradict the implications of the traditional expectations theory. DTSMs neglect frictions in the formation of expectations a possible channel. The interpretation of the pricing kernel in Equation 14.13 is that investors demand compensation for bearing the $\mathcal{N}_{\mathcal{Z}}$ risks captured by $\epsilon_{t+1}^{\mathbb{P}}$. The vector λ_t corresponds to the market prices of risk. Essentially affine models generalize the completely affine models of Duffie and Kan (1996) and Dai and

[13] The last line in 14.15 follows from evaluating the left-hand side at $u = 0$, which requires that

$$1 = E_t^{\mathbb{P}}[\exp(\lambda_t'\epsilon_{t+1}^{\mathbb{Q}})]\exp(-\frac{1}{2}\lambda_t'\lambda_t).$$

Singleton (2000), where the market prices of risk are simply proportional to the yield volatility. The additional reduced-form parameters introduce the potential for overfitting. In addition, DTSMs are typically estimated using yield data only, neglecting the restrictions that prices of risk should reflect investors' preferences and economic fundamentals. The interpretation of λ_t estimates may attribute too much of the risk premium to interest rate risk.

Unspanned Macroeconomic Variables Consider the links between economic variables such as inflation and consumption growth to the observed yields.[14] Duffee (2013) explores the spanning properties of various measures of inflation and consumption growth for yields, and vice versa, and concludes that a large fraction of nominal yield variations is not related to current and past inflation and economic growth. Smoothing the inflation or the consumption data only slightly reduces the cross-sectional deviations relative to unsmoothed data. On the other hand, this reduces the accuracy of inflation and consumption forecasts. Moreover, the risk premia inferred from these affine no-arbitrage models with macroeconomic data are not well estimated and they are inconsistent with evidence from bond returns. Duffee (2013) concludes that one cannot use nominal term structure to obtain reasonable risk compensation for inflation and consumption growth.

Closing the model Feunou and Fontaine (2014) propose one mechanism that directly introduces a wedge between the dynamics for $\mathcal{Z}_t$ and the evolution of the expectations $\mathcal{E}_t$ driving yields. The key feature of the model is that the yield curve must produce efficient forecasts of the state variables, potentially spanning forecasts forecasts of macro variables but not their current values. In the example of inflation, bond yields would span inflation expectations but not inflation itself. The mechanism driving this wedge is distinct from that in Joslin, Priebsch and Singleton (2010), who introduce a class of DTSMs where macro variables are not spanned by the yield curve (i.e., with unspanned macro variables) due restrictions on the prices of risk.

From Equation 14.9, the yield factors $\mathcal{E}_t$ span the conditional expectations of $\mathcal{Z}_{t+1}$ by construction: $\mathcal{Z}_{t+1} = \mathcal{E}_t + u_{t+1}$ for some zero-mean unpredictable innovation u_{t+1}. The key assumption to complete the CM-DTSM specifies the dynamics of $\mathcal{Z}_{t+1}$. We assume that the innovation u_{t+1} corresponds to a rotation of the innovations in $\mathcal{E}_{t+1}$,

$$\mathcal{Z}_{t+1} = \mathcal{E}_t + \Sigma\epsilon^{\mathbb{P}}_{t+1}, \tag{14.21}$$

where Σ is lower diagonal and positive-definite. In other words, we have that $u_{t+1} \equiv \Sigma\epsilon^{\mathbb{P}}_{t+1}$. Combining Equations 14.10 and 14.21 yields

$$\mathcal{E}_{t+1} = K^{\mathbb{P}}_0 + (I_{\mathcal{N}_\mathcal{Z}} + K^{\mathbb{P}}_1)\mathcal{E}_t + \Sigma_{\mathcal{E}}\Sigma^{-1}u_{t+1}, \tag{14.22}$$

[14] Alternatively, one can impose statistical restrictions more than economic ones, replicating the effects of investors' preferences and constraints.

from which we can easily see that $\mathcal{Z}_t$ a VAR(1) if $\mathcal{E}_{t+1} = K_0^{\mathbb{P}} + (I_{\mathcal{N}_{\mathcal{Z}}} + K_1^{\mathbb{P}})\mathcal{Z}_t$. This arises only if

$$\Sigma_{\mathcal{E}}\Sigma^{-1} = I_{\mathcal{N}_{\mathcal{Z}}} + K_1^{\mathbb{P}}, \tag{14.23}$$

with cross-equation restrictions given in Feunou and Fontaine (2014).

Introducing Macro Variables If condition 14.23 holds, then we have that yields must span the forecasts $\mathcal{E}_t$ *and* the state variables $\mathcal{Z}_t$:

$$y_t = a_n + b_n'\mathcal{E}_t = (a_n + b_n'K_0^{\mathbb{P}}) + b_n'(I_{\mathcal{N}_{\mathcal{Z}}} + K_1^{\mathbb{P}})Z_t.$$

Otherwise, yields only span $\mathcal{E}_t$ in the unrestricted case, leaving a separate role for $\mathcal{Z}_t$ to span the information contained in the macro variables. Take a number $0 \le \mathcal{N}_M < \mathcal{N}_{\mathcal{Z}}$ of macro variables $\mathcal{M}_t$ that are spanned by the state $\mathcal{Z}_t$,

$$\mathcal{M}_t = \gamma_0 + \gamma_1\mathcal{Z}_t, \tag{14.24}$$

where γ_0 is a $\mathcal{N}_M \times 1$ vector and γ_1 is a $\mathcal{N}_M \times \mathcal{N}_Z$ rectangular matrix. In addition, take $\mathcal{N}_L = \mathcal{N}_Z - \mathcal{N}_M$ observed portfolios of yields $\mathcal{P}_t$ that are measured without errors,

$$\mathcal{P}_t = Wy_t = a_W + b_W\mathcal{E}_t,$$

where y_t stacks the cross-section of $\mathcal{N} \ge \mathcal{N}_{\mathcal{Z}}$ individual yields, W is a $\mathcal{N}_L \times \mathcal{N}$ matrix of weights, $a_W \equiv W[a_{n_1} \ldots a_{n_{\mathcal{N}}}]'$, and $b_W = W[b_{n_1} \ldots b_{n_{\mathcal{N}}}]'$. This framework can accommodate different numbers of observable macro variables $\mathcal{M}_t$ and yield portfolios $\mathcal{P}_t$. Feunou and Fontaine (2014) show that any canonical conditional mean macro-finance term structure model (CM-MTSM) based on $\mathcal{Z}_t$ is equivalent to a unique canonical CM-MTSM whose $\mathcal{N}_Z \times 1$ state variables are the observable macro variables and yield factors $\mathcal{X}_t' = [\mathcal{M}_t\ \mathcal{P}_t']$ with

$$\begin{aligned}
\mathcal{X}_{t+1} &= \mathcal{E}_{\mathcal{X},t}^{\mathbb{Q}} + \Sigma_X\epsilon_{\mathcal{X},t+1}^{\mathbb{Q}} \\
\mathcal{X}_{t+1} &= \mathcal{E}_{\mathcal{X},t}^{\mathbb{P}} + \Sigma_X\epsilon_{\mathcal{X},t+1}^{\mathbb{P}},
\end{aligned} \tag{14.25}$$

and where the dynamics of $\mathcal{E}_{\mathcal{X},t}^{\mathbb{P}} \equiv E_t[\mathcal{X}_{t+1}]$ under $\mathbb{Q}$ and $\mathbb{P}$ are given by

$$\begin{aligned}
\Delta\mathcal{E}_{\mathcal{X},t+1}^{\mathbb{P}} &= K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}}\mathcal{E}_{\mathcal{X},t}^{\mathbb{P}} + \Sigma_{\mathcal{E}_{\mathcal{X}}}\epsilon_{\mathcal{X},t+1}^{\mathbb{Q}} \\
\Delta\mathcal{E}_{\mathcal{X},t+1}^{\mathbb{P}} &= K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}}\mathcal{E}_{\mathcal{X},t}^{\mathbb{P}} + \Sigma_{\mathcal{E}_{\mathcal{X}}}\epsilon_{\mathcal{X},t+1}^{\mathbb{P}}.
\end{aligned} \tag{14.26}$$

Bond risk premium Importantly, $\mathcal{E}_t$ is not the conditional mean of $\mathcal{Z}_{t+1}$ under the risk-neutral measure $\mathbb{Q}$. This is given by

$$\begin{aligned}
\mathcal{E}_t^{\mathbb{Q}} &\equiv E_t^{\mathbb{Q}}[\mathcal{Z}_{t+1}] \\
&= \mathcal{E}_t + \Sigma E_t^{\mathbb{Q}}[\epsilon_{t+1}^{\mathbb{P}}] \\
&= \mathcal{E}_t + \Sigma\Sigma_{\mathcal{E}}^{-1}E_t^{\mathbb{Q}}[\Delta\mathcal{E}_{t+1} - K_0^{\mathbb{P}} - K_1^{\mathbb{P}}\mathcal{E}_t] \\
&= \mathcal{E}_t + \Sigma\Sigma_{\mathcal{E}}^{-1}\left(K_0^{\mathbb{Q}} - K_1^{\mathbb{Q}}\mathcal{E}_t - K_0^{\mathbb{P}} - K_1^{\mathbb{P}}\mathcal{E}_t\right) \\
&= \mathcal{E}_t + \Sigma\Sigma_{\mathcal{E}}^{-1}\left(K_0^{\mathbb{Q}} - K_0^{\mathbb{P}}\right) + \Sigma\Sigma_{\mathcal{E}}^{-1}\left(K_1^{\mathbb{Q}} - K_1^{\mathbb{P}}\right)\mathcal{E}_t.
\end{aligned} \tag{14.27}$$

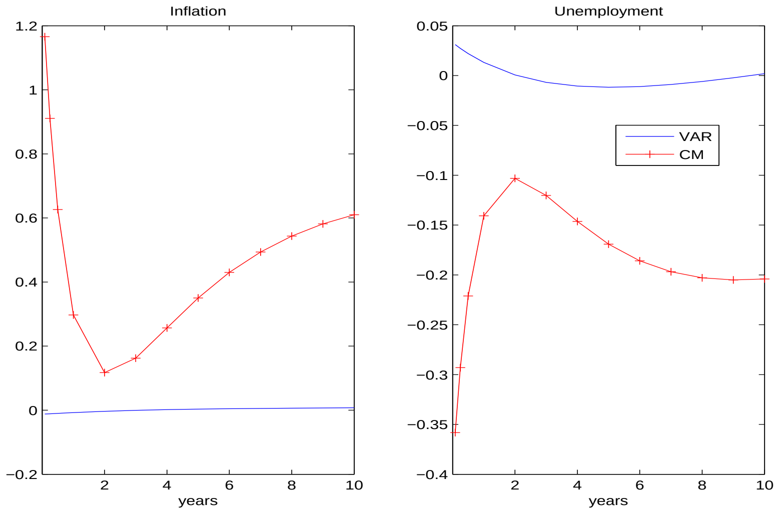


FIGURE 14.10 **Nominal Yield Loadings.**
Loadings of nominal yields on the conditional mean of inflation and on the conditional mean of unemployment from the benchmark CM model.

Therefore, the risk premium $\mathcal{E}_t^{\mathbb{Q}} - \mathcal{E}_t$ is given by a rotation of the volatilities $\Sigma\Sigma_{\mathcal{E}}^{-1}$. However, the risk factors driving the variation of the risk premium satisfies the additional restrictions that they must forecast both the yields and the macro variables $\mathcal{M}_t$.

Results Empirically, Feunou and Fontaine (2014) fit a different version of a CM-MTSM to yields, inflation swaps, surveys of inflation forecasts and unemployment data. The estimates imply that yields span macro expectations in the CM-MTSM. Figure 14.10 reports the loadings on expected inflation and expected unemployment from the unrestricted CM-MTSM and the nested VAR-MTSM . Estimates from the CM-MTSM imply that yield loadings on expected inflation range from 1.16 to 0.61 between the short rate and the 10-year bond, and loadings on expected unemployment range between 0.36 and 0.20. In contrast, these loadings are essentially zero in the VAR. In addition, the CM model produces the best inflation forecasts, supporting the assumptions that yields span $\mathcal{E}_t$. Forecasts from the CM model and forecasts from surveys produce similar out-of-sample root-mean-squared errors (RMSEs) for both Canada and the United States. Relative to the VAR, the CM model reduces the RMSE by as much as 20% when forecasting 2-year cumulative inflation.

14.3.3 SPANNING SURVEY FORECASTS

CM-MTSMs reconcile two strands of the literature. The first uses observed macro variables as risk factors driving yields (inflation, say). The second uses (inflation) survey forecasts Chun (2011). However, in the latter case, shocks to macro variables are left outside the model and the corresponding risk premium (the inflation-risk premium) is unidentified. In contrast, CM-MTSMs can include *both* macro variables and their survey forecasts among the state variables *and* guarantee that statistical forecasts correspond to the survey forecast up to a measurement error.

To see how this works, consider a simple case where inflation π_t and its survey forecast $\bar{\pi}_t^s$ are included in the vector of observables $\mathcal{M}_t$, potentially with other macro variables. By assumption, we have that

$$\pi_t = \gamma_{\pi 0} + \gamma_\pi' \mathcal{Z}_t, \tag{14.28}$$

but also that

$$\mathcal{E}_{\pi,t} = E_t[\pi_{t+1}] = \gamma_{\pi 0} + \gamma_\pi' E_t[\mathcal{Z}_{t+1}] = \gamma_{\pi 0} + \gamma_\pi' \mathcal{E}_t. \tag{14.29}$$

Suppose that $\bar{\pi}_t^s$ and $\mathcal{E}_{\pi,t}$ differ by some measurement errors $\eta_{\pi,t}^s$,

$$\bar{\pi}_t^s = \gamma_{\pi 0} + \gamma_\pi' \mathcal{E}_t + \eta_{\pi,t}^s. \tag{14.30}$$

With no loss of generality, consider the case where π_t and $\bar{\pi}_t^s$ are ordered first and second in the vector $\mathcal{X}_t$, respectively.[15] Then, the following conditions

$$e_1' K_{0X}^{\mathbb{P}} = 0 \quad \text{and} \quad \delta \equiv e_2 - (I + K_{1X}^{\mathbb{P}})' e_1 = 0$$

guarantee that (i) $E^{\mathbb{P}}[\eta_{\pi,t}^s] = 0$ and (ii) $\text{cov}(\eta_{\pi,t}^s, \eta_{\pi,t+h}^s) = 0 \quad h > 0$ so that $\eta_{\pi,t}^s$ is a measurement error (e_i has its i-th element equal to 1 and equal to 0 otherwise). In other words, $e_1' K_{0X}^{\mathbb{P}} = 0$ requires that the first element of the constant be zero and $\delta = 0$ requires that the second line of the autoregressive matrix be equal to e_2'. Clearly, these conditions can be implemented easily. They also have an intuitive economic interpretation:

$$\begin{aligned}
\mathcal{E}_{\pi,t} &= e_1' \mathcal{E}_{\mathcal{X},t} = e_1' (K_{0X}^{\mathbb{P}} + (I + K_{1X}^{\mathbb{P}}) \mathcal{E}_{\mathcal{X},t-1} + \Sigma_{\mathcal{E}_\mathcal{X}} \epsilon_t^{\mathbb{P}}) \\
&= e_2' \mathcal{E}_{\mathcal{X},t-1} + e_1' \Sigma_{\mathcal{E}_\mathcal{X}} \epsilon_t^{\mathbb{P}} \\
&= e_2' \mathcal{X}_t + (e_1' \Sigma_{\mathcal{E}_\mathcal{X}} - e_2' \Sigma_X) \epsilon_t^{\mathbb{P}} \\
&= \bar{\pi}_t^s - e_1' \theta_X^{\mathbb{P}} \Sigma_X \epsilon_t^{\mathbb{P}},
\end{aligned} \tag{14.31}$$

where $\theta_X^{\mathbb{P}} \equiv I + K_{1X}^{\mathbb{P}} - \Sigma_{\mathcal{E}_\mathcal{X}} \Sigma_X^{-1}$. The model-implied forecast corresponds to the survey forecasts plus a rotation of innovations in $\mathcal{X}_t$ (including the innovations in survey forecasts). This guarantees that Equation (14.30) holds.[16]

[15] Feunou and Fontaine (2014) discuss the general case with several macro variables included, along with their corresponding survey forecasts.

[16] The necessary and sufficient conditions for $\text{cov}(\eta_t^s, \eta_{t+h}^s) = 0 \quad h > 0$ are given in Appendix A6 of Feunou and Fontaine (2014). The sufficient conditions introduced above impose no restrictions on the covariance matrix (which are not free in the canonical form).

Equation 14.31 also shows that Markovian VAR(1) models (i.e., if $\theta_X^{\mathbb{P}} = 0$) can accommodate Equation (14.30) but only at the added cost that $\text{var}(\eta_t^s) = 0$. In other words, the inflation forecasts from the model must correspond *exactly* to the survey forecasts, again leading to the counterfactual prediction that yields span the survey forecasts exactly.

Subjective Expectations So far, we assumed that investors' historical predictions were identical to predictions produced by statistical models estimated in-sample today up to a measurement error. However, the evidence discussed in Section 14.3.1 forces the question as to wether the assumption of rational expectations is an appropriate assumption in the context of term structure models.

As in Cieslak and Povala (2013), Piazzesi, Salomao and Schneider (2013) conclude that subjective interest rate expectations deviate from the expectations computed from statistical models and these deviations seem to explain most of the statistical bond premia. They then proceed to study the relative contribution of beliefs and risk assessment for bond return predictability within a no-arbitrage model. They introduce a subjective change of measure causing a wedge between forecasts from the statistical model and from surveys (see also Chernov and Mueller, 2011). Introducing a subjective probability measure provides allows a more general approach to model the dynamics of survey forecasts relative to including the survey forecasts among the state variables.

We can use the notation introduced above while staying close to the approach in Piazzesi, Salomao and Schneider (2013). Suppose the following subjective change of measure:

$$\xi_{t+1}^* \equiv \frac{\exp(\lambda_t^{*'} \epsilon_{t+1}^{\mathbb{Q}})}{E_t^*[\exp(\lambda_t^{*'} \epsilon_{t+1}^{\mathbb{Q}})]}, \tag{14.32}$$

and the following subjective prices of risk:

$$\lambda_t^* \equiv \lambda_0^* + \lambda_1^{*'} \mathcal{E}_t. \tag{14.33}$$

Then, following the same argument as in Equation 14.15, it follows that the dynamics of $\mathcal{E}_t$ under the subjective measure are given by

$$\Delta\mathcal{E}_{t+1} = K_0^* + K_1^* \mathcal{E}_t + \Sigma_{\mathcal{E}} \epsilon_{t+1}^*. \tag{14.34}$$

with parameters given by

$$K_0^* = K_0^{\mathbb{Q}} - \Sigma_{\mathcal{E}} \lambda_0^* \tag{14.35}$$

$$K_1^* = K_1^{\mathbb{Q}} - \Sigma_{\mathcal{E}} \lambda_1^*. \tag{14.36}$$

Equation14.34 provides the dynamics of the state variables under the subjective measure and yields a set of measurement equations corresponding to observable survey forecasts. We can write

$$\bar{X}_t^s = E_t^*[X_{t+1}] + \eta_t^s = K_0^* + (I + K_1^*)\mathcal{E}_t + \eta_t^s \tag{14.37}$$

where $\bar{X}_t^s$ is the observed survey forecast and η_t^s is a measurement error.

Piazzesi, Salomao and Schneider (2013) implement a 3-factor model that includes a change of measure due to risk, as in Equation 14.13, as well as a change of measure due to subjective beliefs, as in Equation 14.32. As observable state variables, Piazzesi, Salomao and Schneider (2013) use the level of yields, the slope of the yield curve and expected inflation. They estimate each set of parameters sequentially. First, they estimate the historical dynamics of the state vector by maximum likelihood. Second, they estimate parameters of the risk-neutral measure using observed yields. Together with the historical dynamics, Piazzesi, Salomao and Schneider (2013) can then compute estimates of the "statistical" risk premium. Finally, they estimate parameters of the subjective measure using survey forecasts. Together with the risk-neutral dynamics, they can then compute estimates of the subjective risk premium.

Piazzesi, Salomao and Schneider (2013) find that the level and slope of the yield curve are less persistent under the subjective measure—estimated from surveys—than under the statistical model. Hence, the subjective risk premium is much less cyclical than the statistical premium. In the statistical models, a high level of yields and a high slope in the yield curve predict higher excess bond returns. Survey forecasters, who view the level and the slope as more persistent, predict lower excess bond returns. Overall, this reduces the volatility of subjective risk premia and lengthens the frequency of its variations.

Garcia and Luger (2012) also explore investors' perceived factor dynamics but in a framework with recursive utility and a time-varying subjective discount parameter.[17] They use VAR(1) dynamics for the state variables $Z_t = (y_t^{(1)}, y_t^{(20)} - y_t^{(1)}, r_t, \pi_t, c_t)')$, where the first two variables are the one-quarter and the 5-year spread, r_t is the log-return on the wealth portfolio, π_t the inflation rate and c_t the log-consumption growth. The Euler equation defines the logarithm of the nominal stochastic discount factor:

$$\log m_{t+1}^{\$} = -y_t^{(1)} + J'\mathcal{Z}_{t+1}, \tag{14.38}$$

where $J = (0, 0, \gamma - 1, -1, \alpha - \gamma)'$. The parameters γ and α determine the coefficients of intertemporal substitution and risk aversion, respectively. Bond prices are parameterized as exponential linear functions of the state vector so that

$$P(t, n) = \exp\left(A(n) + B(n)'\mathcal{Z}_t\right), \tag{14.39}$$

with coefficients $A(n)$ and $B(n)$ determined by the backward recursions

$$\begin{aligned} A(n+1) &= A(n) + [J + B(n)]'\mu + \tfrac{1}{2}[J + B(n)]'\Sigma\Sigma'[J + B(n)], \\ B(n+1) &= \Phi'[J + B(n)] - e_1, \end{aligned} \tag{14.40}$$

and these also depend on the parameters of the VAR(1) dynamics (with $e_1' = (1, 0, 0, 0, 0)$).

[17]Following Obstfeld (1990), the subjective discount parameter is linked to the short-term rate and becomes a state variable in their framework.

Since the compensation for risk is driven by only two parameters, Garcia and Luger (2012)note that the observation equation for yields identifies all the parameters of the historical dynamics. Hence, they proceed to estimate all the parameters directly from market yields. The recovered dynamics must be consistent with the expectations of a representative investor with recursive utility that is pricing observed market yields.

Estimates of the preference parameters are economically plausible: the coefficient of relative risk aversion is around 6 and the elasticity of intertemporal substitution is around 0.36. The estimates for μ, Φ and Σ correspond to the *historical* dynamics of Z_t. These produce bond yields and returns that match empirical stylized facts. In regressions of future excess returns on forward rates, as in Cochrane and Piazzesi (2005), the recursive utility model produces slope coefficients with the correct pattern across maturities, and the confidence intervals cover most of the coefficients estimated via direct OLS in market data. Furthermore, the hump-shaped pattern in the term structure of unconditional volatilities of yields and yield changes is apparent.

The perceived dynamics implicit in bond yields differ from the estimates obtained by an econometrician using the available data for Z_t. To illustrate this wedge, they transpose the AR coefficients estimated directly from the bond yields but using the equilibrium model into a standard reduced-form no-arbitrage version of the model. This modified reduced-form model then becomes competitive with the recursive utility model in reproducing empirical stylized facts for bond returns, but it does not provide any improvement over the equilibrium approach, despite the fact that 30 prices of risk are estimated to fit bond yields.

14.3.4 ADAPTIVE LEARNING AND SURVEY FORECASTS

One reading of the evidence— following Kim and Orphinides (2012), say—is that survey data should be used as additional information to estimate the dynamics of yield factors (as perceived by bond investors). An alternative reading, however, is that surveys do not offer a reliable basis to estimate the historical dynamics. But is it plausible that bond prices carry the same expectation "errors" as surveys? In a sense, this brings us back to our original conundrum.

One approach to this question is to consider that agents observe the relevant state variables, but that they are learning about unknown parameters. A first question to ask is whether the adaptive learning approach is able to replicate the expectations of professional forecasters in survey data. A recent paper by Markiewicz and Pick (2014) argues that learning models provide a good fit of survey forecasts for a range of macroeconomic and financial variables. They consider simple time-series autoregressive models to make predictions of the survey forecasts and update the parameters—generically $y_t = \boldsymbol{\theta}_t' \boldsymbol{x}_t + \varepsilon_t$)—using recursive least squares, as follows:

$$
\begin{aligned}
\boldsymbol{\theta}_t &= \boldsymbol{\theta}_{t-1} + \gamma_t \mathbf{R}_t^{-1} \boldsymbol{x}_t (y_t - \boldsymbol{\theta}_{t-1}' \boldsymbol{x}_t), t = 1, 2, \ldots, T, T+1, \\
\mathbf{R}_t &= \mathbf{R}_{t-1} + \gamma_t (\boldsymbol{x}_t \boldsymbol{x}_t' - \mathbf{R}_{t-1}), \qquad (14.41)
\end{aligned}
$$

where $\boldsymbol{\theta}_t$ is a $k \times 1$ vector of parameters and $\boldsymbol{x}_t$ is the $k \times 1$ vector of regressors, possibly including lags, and γ_t is the gain parameter. Decreasing gain learning uses $1/t$ and corresponds to standard OLS with a widening window. Constant gain learning set γ to a constant that is estimated in a second step by minimizing the distance with survey forecasts. Markiewicz and Pick (2014) find that constant gain learning provides a much closer fit to the survey forecasts of inflation, the 3-month bill rate and the 10-year bond yield.

A second question to ask is whether allowing for learning can rationalize the evidence against the expectations hypothesis based on predictability regressions of bond returns. Bulkley and Giordani (2011) provide one positive answer. They assume that the short-rate process is subject to repeated stochastic structural breaks, in which the break points are not observed but have to be inferred in real time from short-rate data. Agents must judge whether an interest rate surprise denotes a break and they must estimate new parameter values conditional on that inference. In this context, Bayesian learning induces changes in rational expectations—hence changes in long yields—that are consistent with the evidence recovered in Campbell-Shiller regressions. Their calibration is based on estimates from fitting a model of stochastic breaks to US short rate data.

14.3.5 EQUILIBRIUM MODELS OF THE TERM STRUCTURE

Duffee (2013) reviews several dynamic stochastic general equilibrium (DSGE) models that have been proposed to model the term structure of interest rates. He notes that expected bond returns appear countercyclical: the correlations between bond returns with consumption or industrial growth are negative. Therefore, the workhorse representative agent model with power utility cannot generate a positive average slope, since nominal bonds appear to be a hedge. Allowing for recursive preferences, and considering the case where the elasticity of intertemporal substitution is one, the real stochastic discount factor is given (in log) by

$$m_{t+1} = \log \beta - \Delta c_{t+1} - (\gamma - 1)\zeta_{t+1} - \frac{\gamma - 1}{2}^2 \mathrm{Var}_t(\zeta_{t+1}), \tag{14.42}$$

where β is the discount rate, Δc_{t+1} is consumption growth, γ is the coefficient of risk aversion and ζ_{t+1} is the innovation to expectations of discounted consumption growth:

$$\zeta_{t+1} = (E_{t+1} - E_t) \sum_{i=1}^{\infty} \beta^i \Delta c_{t+1}. \tag{14.43}$$

Nominal bond returns have negative covariance with consumption growth. Therefore, nominal bonds carry a negative risk premium in the benchmark case with power utility ($\gamma = 1$). With recursive utility, the covariance with innovation to future consumption growth also plays a role. For instance, if inflation shocks are positively correlated with lower expected consumption in the future (stagflation), bond returns may be risky if the positive covariance with future consumption growth is high enough. Based on a variety of statistical models estimated in different sample periods, Duffee (2013) concludes that there is not enough information in the estimated joint dynamics of consumption growth and inflation to support the stagflation interpretation of risk premia.

A recent line of research introduced extensions to existing equilibrium models with recursive utility. However, equilibrium models that include elements of learning also appear promising. It turns out that the dynamics of the pricing kernel can differ considerably from the dynamics of realized consumption, because of subjective beliefs. Collin-Dufresne, Johannes, and Lochstoer (2013) study a long-risk economy with recursive utility similar to Bansal and Yaron (2004).[18] They show that parameter learning strongly amplifies the impact of macro shocks on marginal utility when the representative agent has a preference for early resolution of uncertainty. This occurs as rational belief updating generates subjective long-run consumption risks. For instance, even a small update to the mean of consumption growth can have a large effect on the continuation utility, since the shock will affect the expected consumption growth far into the future. The authors derive the term structure in a setting with an intertemporal elasticity of substitution equal to 1. With similar parameter settings as in Bansal and Yaron (2004), they obtain a flat yield curve, which is quite different from the strongly downward-sloping curve obtained by Bansal and Yaron (2004). The difference derives from the fact that updates in beliefs are permanent, as opposed to the mean-reverting process assumed by Bansal and Yaron (2004) for mean consumption growth.

Looking beyond asset pricing, Del Negro and Eusepi (2011) investigate whether DSGE models estimated under rational expectations or models with learning about target inflation are a better fit with inflation forecasts from surveys. Ormeno (2009) also estimates learning models with data on inflation expectations extracted from surveys. Milani (2007) estimates a small-scale DSGE model with learning using only historical data on macroeconomic variables and not survey forecast data. He shows that adaptive learning enhances the propagation mechanism and reduces the scale of rational frictions and the persistence that they cause. Slobodyan and Wouters (2012) extend their work in a medium-scale DSGE model. Their results illustrate that learning models fit the data as well as, or better than, the rational-expectations model.[19]

14.4 Lower Bound

Currency yields no interest but coexists with interest-bearing bonds. The risk-free bond may be preferable as a store of value (it yields a higher return), but

[18] Early examples where asset pricing implication are driven by learning in a general equilibrium model include Detemple (1986); Dothan and Feldman (1986).

[19] However, their findings are crucially dependent on the assumptions made about initial beliefs. The best models correspond to the case where initial beliefs are optimized to explain the in-sample data. In this case, the updating of beliefs through constant-gain learning brings only a low estimated gain.

holding currency may nonetheless be desirable for its use as a medium of exchange. With no frictions, currency becomes a better store of value if the interest rate on the risk-free bond becomes negative, $r_t < 0$. In this case, investors should issue, sell or short the bond; hold currency; and expect a sure profit. This arbitrage keeps the interest rate above zero.

With transaction costs or short-selling frictions, bond yields remain bounded but not necessarily at zero, which explains the choice of title for this section. If holding currency is costly—banks charge a fee to store currency—then the bond could carry a negative yield. On the other hand, if holding the safe bond is costly—the money-market mutual funds charge a fee— then the bond yield may be bounded above zero. The Federal Reserve banks and several other central banks refrained from pinning down their target rates to zero. More recently, the European Cental Bank, the Swiss National Bank and others lowered their target rate below zero. In any case, the short rate can be written as

$$r_t = lb_t + r_t^*, \tag{14.44}$$

where lb_t is the lower bound for the nominal short rate and r_t^* is a positive process. If lb_t is stochastic, it becomes one of the factors driving the observed short rate. In the following, we take $lb_t = 0$ for simplicity.

14.4.1 SQUARE-ROOT AND AUTO-REGRESSIVE GAMMA MODELS

Consider the standard affine specification of the short rate:

$$r_t = \delta_0 + \delta_1' x_t, \tag{14.45}$$

where x_t includes K risk factors. For $\delta_0 \geq 0$ and $\delta_1 \geq 0$, the short rate remains positive whenever each risk factor remains positive. Cox, Ingersoll and Ross (1985) introduced the continuous-time square-root process to the term structure literature. In this case, each $x_{i,t}$ for $i = 1, \dots, K$ follows an independent square-root process:

$$dx_{i,t} = \kappa_i(\theta_i - x_{i,t}) + \sigma_i \sqrt{x_{i,t}} dW_{i,t}, \tag{14.46}$$

with parameters κ_i, θ_i and σ_i. This process is strictly positive when Feller's condition is satisfied: $\frac{2\kappa\theta}{\sigma^2} \geq 1$. Otherwise, zero is an absorbing state: the process may reach zero, but if it does it cannot leave zero.

The square-root process has a discrete-time counterpart: the autoregressive gamma (ARG) process. Let us detail how the ARG process is constructed. Its basic ingredients are the Poisson distribution $\mathcal{P}(\lambda)$ with intensity $\lambda > 0$ and the standard gamma distribution $\gamma(\nu)$ with shape $\nu > 0$. In the univariate case, following Gourieroux and Jasiak (2006) and Le, Singleton and Dai (2010), the conditional distribution of $x_{i,t}$ is defined as a Poisson mixture of standard gamma distributions:

$$x_{i,t} | x_{i,t-1} \sim \gamma(\nu_i + W_i) \quad \text{where} \quad W_i \sim \mathcal{P}(\rho_i x_{i,t-1}/c_i), \tag{14.47}$$

with parameter $c_i > 0$, $\nu_i > 0$ and $0 \leq \rho_i < 1$. The conditional mean of $x_{i,t}$ is autoregressive:

$$x_{i,t} = c_i \nu_i + \rho_i x_{i,t-1} + \sigma_{i,t} \epsilon_{i,t}, \tag{14.48}$$

where $\epsilon_{i,t}$ has mean zero, $E_t[\epsilon_{i,t}] = 0$ and with conditional variance given by

$$\sigma_{i,t}^2 = c_i^2 \nu_i + 2 c_i \rho_i \, x_{i,t-1}. \tag{14.49}$$

The conditional Laplace transform of $x_{i,t}$ is exponential-affine (see Gourieroux and Jasiak, 2006):

$$E_{t-1}[e^{u x_{i,t}} | x_{i,t-1}] = \exp\left(-\nu_i \ln(1 - c_i u) + \frac{u \rho_i x_{i,t-1}}{1 - c_i u} \right), \quad u < 1/c_i. \tag{14.50}$$

If $\nu > 1$, then the limit of (14.48) as the time interval Δt shrinks to zero is a square-root process. Conversely, when the continuous-time process satisfies Feller's condition, $\frac{2\kappa\theta}{\sigma^2} \geq 1$, then the distribution given by (14.46) when sampling at frequency Δt corresponds exactly to the distribution given by (14.50). The exact mapping between parameters is given by

$$\rho = e^{-\kappa \Delta t} \qquad c = (1 - \rho)\frac{\sigma^2}{2\kappa} \qquad \nu = \frac{2\kappa\theta}{\sigma^2} = (1 - \rho)\frac{\theta}{c}, \tag{14.51}$$

which also makes clear the connection between the Feller condition and the discrete-time parameter ν.

Reaching the lower bound If $\nu < 1$, then zero is an absorbing state for the continuous-time process. However, it should be clear from (14.47) that zero is not an absorbing state for the ARG process even if $\nu < 1$. In fact, as long as $\nu > 0$, this process never reaches zero. However, there is no equivalence with the continuous-time case when $\nu < 1$.

Monfort et al. (2014) build the ARG_0 process, allowing $x_{i,t}$ to reach zero, stay at zero for extended periods of time, and eventually leave zero. This is achieved by setting $\nu_i = 0$ but introducing a constant α_i in the intensity of the Poisson distribution (compare with Equation 14.47):

$$x_{i,t} | x_{i,t-1} \sim \gamma(W_i) \quad \text{where} \quad W_i \sim \mathcal{P}(\alpha_i + \rho_i' x_{i,t-1}/c_i), \tag{14.52}$$

for the multivariate case ($\rho_i > 0$ is a vector), recognizing that $\gamma(\nu)$ converges to the Dirac distribution when ν reaches zero. Heuristically, the mixing variable W_i is zero with strictly positive probability, in which case the realization of $x_{i,t}$ is also zero (almost surely). The conditional Laplace transform of the ARG_0 process differs substantially from (14.50):

$$E_{t-1}[e^{u x_{i,t}} | x_{i,t-1}] = \exp\left(\frac{u(\alpha_i c + \rho_i' x_{i,t-1})}{1 - c_i u} \right), \quad u < 1/c_i. \tag{14.53}$$

The parameter α_i quantifies the average persistence of zero lower-bound regimes. First, the probability that $x_{i,t}$ reaches zero h periods from now is

$$P(x_{i,t} = 0 | x_{i,t-1}) = \exp(-h\alpha_i - \rho_i' x_{i,t-1}/c_i). \tag{14.54}$$

Second, denote $p = P(x_{i,t} = 0|x_{i,t-1} = 0)$, the probability of staying at zero. Fixing $h = 1$ and $x_{i,t-1} = 0$, it follows that $p = e^{-\alpha_i}$. Of course, the probability of leaving zero is $P(x_{i,t} > 0|x_{i,t-1} = 0) = 1 - p$. Third, the probability of leaving zero after $h - 1$ periods at zero is

$$P(x_{i,t} = 0, \ldots, x_{i,t+h-1} = 0, x_{i,t+h} > 0|x_{i,t-1} = 0) = (1-p)p^{h-1}, \quad (14.55)$$

and, finally, the average sojourn time at zero if $x_{i,t} = 0$ is given by

$$\frac{1}{1-p} = \left(1 - e^{-\alpha_i}\right)^{-1}. \quad (14.56)$$

Term structure models Monfort et al. (2014) mix ARG and ARG_0 processes to construct a new family of term structure models with positive interest rates where the short-rate can stay at its lower bound for long periods of time. This is a significant step forward relative to the standard square-root or ARG cases. On the other hand, models based on ARG_0 processes retains some notable weaknesses: (i) it is hard to imagine using (possibly) negative macro variables extensively in this context, (ii) the correlation structure between risk factors is severely limited, since every correlation must be non-negative and (iii) the admissibility constraint gives rise to a tension in simultaneously fitting of the physical and risk-neutral yields forecasts (Joslin and Le (2011)).

The short rate process is given by

$$r_t = \delta_1' x_t^{(1)} + \delta_2' x_t^{(2)} = \delta' x_t, \quad (14.57)$$

where $x^{(1)}$ includes n_1 components, $x_t^{(2)}$ includes n_2 components (with $n = n_1 + n_2$), and with $x_t' = [x^{(1)'} \; x^{(2)'}]$. The dynamics of each component $x_{i,t}$ under the historical probability measure $\mathbb{P}$ is characterized by its Laplace transform $\varphi_{i,t}(u) \equiv E_{t-1}[e^{ux_{i,t}}|x_{i,t-1}]$:

$$\varphi_{i,t}(u) = \exp\left(-\nu_i \ln(1 - c_i u) + \frac{u(\alpha_i c_i + \rho_i' x_{t-1})}{1 - c_i u}\right) \quad u < 1/c_i, \quad (14.58)$$

while the asset pricing implications follow from the corresponding transform $\varphi_{i,t}^{\mathbb{Q}}(u)$ under the risk-neutral measure $\mathbb{P}$:

$$\varphi_{i,t}^{\mathbb{Q}}(u) = \exp\left(-\nu_i \ln(1 - c_i^{\mathbb{Q}} u) + \frac{u(\alpha_i^{\mathbb{Q}} c_i^{\mathbb{Q}} + \rho_i^{\mathbb{Q}'} x_{t-1})}{1 - c_i^{\mathbb{Q}} u}\right) \quad u < 1/c_i, \quad (14.59)$$

which is similar to the multivariate extension of the ARG proces in Le, Singleton and Dai (2010). The novel feature is the parameter α_i and the allowance for $\nu_i \geq 0$. In addition to $c_i > 0$ and $\rho > 0$ (element-by-element), their family of models has that

1. $0 <, \rho_{ii}^{\mathbb{P}}, \rho_{ii}^{\mathbb{Q}} < 1, \; i = 1 \ldots n$

2. $c_i^{\mathbb{Q}} = 1, \; i = 1 \ldots n$

3. $\nu_i^{\mathbb{P}} = \nu_i^{\mathbb{Q}} = 0,\ i = 1 \ldots n_1$

4. $\delta_i^{(2)} = 0,\ i = 1 \ldots n_2.$

Assumption 1 implies that x_t is stationary. Assumption 2 is required for identification. More importantly, Assumptions 3 imply that the first n_1 factors are ARG_0 processes. Note that the absence of arbitrage requires that ν_i cannot be zero under the historical measure and different zero under the risk-neutral measure.[20] Assumption 4 implies that the short rate is driven by the n_1 ARG_0 factors. Together, these two assumptions imply that the short rate has a probability mass at zero.

Under Assumptions 1–4, the solution for bond prices $P_t^{(n)}$ is exponential-affine $P_t^{(n)} = \exp\left(A_n + B_n' \mathcal{E}_t\right)$, as in the previous section, but with coefficients given by

$$\begin{aligned} B_{n+1} &= -\delta + \rho^{\mathbb{Q}'} \left(\frac{B_n \odot \iota}{1 - A_n \odot \iota} \right) \\ A_{n+1} &= A_n + \alpha^{\mathbb{Q}'} \left(\frac{B_n \odot \iota}{1 - A_n \odot \iota} \right) - \nu' \log\left(1 - A_n \odot \iota\right), \end{aligned} \tag{14.60}$$

where, following Monfort et al. (2014), $\odot$ denotes the Hadamard product while the division and $\log$ operators apply to the vector element-by-element (see Proposition 3.2 in Monfort et al. (2014) for closed-form expressions for bond prices).

14.4.2 BLACK (1995) – TOBIT

Black (1995) builds on the observation that money provides an option to the bondholder: the bondholder can choose to hold money if the bond yield becomes negative. From a pure time-series perspective, this problem corresponds to an environment with censored observations: a "shadow" short-rate process may become negative, but the corresponding observations are censored Tobin (1958). Economically, the short rate corresponds to the combination of a (possibly negative) shadow rate s_t and the currency option:

$$r_t = \max(0, \delta_0 + \delta' x_t), \tag{14.61}$$

where the shadow rate is linear $s_t = \delta_0 + \delta' x_t$. An important perceived advantage is that x_t can have Gaussian dynamics. As Dai and Singleton (2002) show, Gaussian models provide a better fit for the risk premium than models with squared-root (i.e., ARG) dynamics when interest rates are far from zero. Arguably, the gains in fitting bond excess returns were considered worth the small probability that yields became negative when the short rate was far from zero. This trade-off is no longer palatable: a Gaussian short rate near zero has nearly 50% of its probability mass below zero.

[20] The probability measures must be equivalent.

It was long thought that asset pricing implications could be hard to derive with a $\max$ function for the short-rate process. Specifically, even if the risk factor dynamics are tractable,

$$\Delta x_t = K_0^{\mathbb{Q}} + K_1^{\mathbb{Q}} x_{t-1} + \Sigma \epsilon_t^{\mathbb{Q}} \quad \epsilon_t^{\mathbb{Q}} \sim N(0, I), \tag{14.62}$$

the computation of bond yields via the Laplace transform,

$$B_t^{(n)} = E_t^{\mathbb{Q}} \left[\exp \left(- \sum_{i=0}^{n} \max(0, s_{t+i}) \right) \right], \tag{14.63}$$

is no longer available. Ichiue and Ueno (2007) and Kim and Singleton (2012) compute the integral (the expectation) via a discretization scheme. Alternatively, Bauer and Rudebusch (2013) proceed via Monte Carlo simulations. Kripnner (2013) provides a broader survey of these numerical approaches.

In the following, we detail a second-order approximation based on the series expansion of 14.63:

$$B_t^{(n)} \approx 1 - E_t^{\mathbb{Q}} \left[\sum_{i=0}^{n} r_{t+i} \right] + \frac{1}{2} E_t^{\mathbb{Q}} \left[(\sum_{i=0}^{n} r_{t+i})^2 \right]. \tag{14.64}$$

Priebsch (2013), in continuous time, and Wu and Xia (2013), in discrete time, follow a similar route based on the series expansion of $\log B_t^{(n)}$. We follow a different route to exploit the recursive structure of a discrete-time model. Krippner (2012) introduces an alternative the representation of (14.63) using the price of a call option. This approximation appears to be slightly less accurate in practice.

Rearranging, we obtain the following recursion for bond prices:

$$\begin{aligned} B_t^{(n+1)} = B_t^{(n)} - & E_t^{\mathbb{Q}} \left[\max(0, s_{t+n}) \right] + E_t^{\mathbb{Q}} \left[\max(0, s_{t+n})^2 \right] \\ & + \sum_{i=0}^{n-1} E_t^{\mathbb{Q}} \left[\max(0, s_{t+i}) \max(0, s_{t+n}) \right], \end{aligned} \tag{14.65}$$

for $n = 1, 2, \ldots$, with $B_t^{(0)} = 0$. To compute bond prices, each term in (14.65) can be computed explicitly using moments of a truncated bivariate normal distribution (Priebsch, 2013; Rosenbaum, 1961). Note that

$$\begin{pmatrix} s_{t+i} \\ s_{t+n} \end{pmatrix} | x_t \sim N^{\mathbb{Q}} \left[\begin{pmatrix} \mu_{i|t} \\ \mu_{n|t} \end{pmatrix}, \begin{pmatrix} \sigma_{i|t}^2 & \sigma_{i,n|t} \\ \sigma_{i,n|t} & \sigma_{n|t}^2 \end{pmatrix} \right] \quad 0 \leq i < n, \tag{14.66}$$

where the mean and covariance terms are once again given by straightforward recursions:

$$\begin{aligned} \mu_{i+1|t} &= \mu_{i|t} + \delta'(I + K_1^{\mathbb{Q}})^i \left(K_0^{\mathbb{Q}} + K_1^{\mathbb{Q}} X_t \right) \\ \sigma_{i+1|t}^2 &= \sigma_{i|t}^2 + \delta'(I + K_1^{\mathbb{Q}})^i \Sigma\Sigma'(I + K_1^{\mathbb{Q}})^{i'} \delta \\ \sigma_{i+1,n|t} &= \sigma_{i,n|t} + \delta'(I + K_1^{\mathbb{Q}}) \Sigma\Sigma'(I + K_1^{\mathbb{Q}})^{n-1-i'} \delta, \end{aligned} \tag{14.67}$$

for $n = 1, 2, \ldots$. The initial conditions are $\mu_{0|t} = \delta_0 + \delta' x_t$, $\sigma^2_{0|t} = 0$, and $\sigma_{0,n|t} = 0$. Then, the bond prices in (14.65) are given explicitly in terms $\mu_{i+1|t}$, $\sigma^2_{i+1|t}$, and $\sigma_{i+1,n|t}$, and in terms of the Gaussian cumulative and probability distribution functions (see the appendix).

This approximation is exact if r_{t+i} is Gaussian (i.e., if $r_{t+i} = \delta_0 + \delta' x_{t+i}$), but higher-order terms come into play once we introduce the max function. Nonetheless, simulation evidence suggests that these terms are small when x_{t+i} is Gaussian (Priebsch, 2013; Wu and Xia, 2013). On the other hand, if x_{t+i} is not Gaussian, the computation of the truncated moments becomes intractable *and* the higher-order terms may not be negligible.

14.4.3 NO-DOMINANCE TERM STRUCTURE MODELS

The limitations and challenges of no-arbitrage models arise because of the desire for tractability and the difficulty of computing expectations, e.g., in (14.63). Exploiting the first fundamental theorem of asset pricing is a powerful tool to guarantee the absence of arbitrage among a set of model prices. Nonetheless, the requirement that yields remain positive limits the choice of dynamic specifications for which we can obtain analytical solutions. Of course, simulation is an acceptable way to compute (14.63), but the benefits of closed forms are unambiguous.

Fontaine, Feunou and Le (2014) circumvent the computation of expectations and the need to specify the risk-neutral measure $\mathbb{Q}$ for bond pricing. Instead, they directly specify bond prices in terms of primitive functionals:

$$B_t^{(n)} \equiv 1, \tag{14.68}$$

$$B_t^{(n)} = B_t^{(n)}(g(X_t)) \times \exp(-m(X_t)), \tag{14.69}$$

for arbitrary functions $m(\cdot)$ and $g(\cdot)$ such that $g(x_t) \in \underline{\mathrm{X}}$ for every $x_t \in \underline{\mathrm{X}}$ the support of the risk factors. They show that bond prices given by Equations 14.68 and 14.69 are free of dominant trading strategies. A dominant trading strategy exists if a zero-cost portfolio has strictly positive payoffs. This portfolio's price,

$$\sum_n w_n B_t^{(n)} = \exp(-m(x_t)) \times \sum_n w_n B_t^{(n)}(g(x_t)), \tag{14.70}$$

is strictly positive since its payoff is strictly positive ($g(x_t) \in \underline{\mathrm{X}}$), hence forbidding dominant strategies. No-dominance still allows for zero-cost portfolios that have zero or positive payoffs. These are knife-edge arbitrage opportunities that, arguably, should be disregarded at *some* level of transaction costs. In fact, Fontaine, Feunou and Le (2014) show that these arbitrage opportunities are precluded in the model as soon as transactions are not zero, either in establishing or selling the portfolio.

Example 1 – Affine short-rate and Nelson-Siegel models From (14.69), we have that $B_t^{(1)} = \exp(-m(x_t))$, implying that $r_t = m(x_t)$. For instance, the choice of $m(x_t) = \delta_0 + \delta' x_t$ and $g(x_t) = K x_t$ leads to the following bond yields:

$$y_t^{(n)} = \delta_0 + (b_n/n) x_t, \tag{14.71}$$

with b_n given by the recursion:

$$b_n = b_{n-1} K + \delta', \tag{14.72}$$

with $b_0 = 0$. The risk-factor loadings in (14.72) are identical to results for the canonical term structure models in Joslin, Singleton and Zhou (2011). However, the constant in the no-arbitrage model,

$$a_n = a_{n-1} + \delta_0 - \frac{1}{2} b_{n-1} \Sigma b'_{n-1}, \tag{14.73}$$

differs from the constant in the no-dominance specification, due to the (small) Jensen's term. Indeed, Fontaine, Feunou and Le (2014) show that 3-factor affine no-dominance and no-arbitrage models exhibit no visible pricing differences. Interestingly, restricting the matrix K as in Christensen, Diebold and Rudebusch (2011) produces the Nelson-Siegel (NS) loadings. In fact, the continuous-time limit corresponds squarely to the adaptation by Diebold and Li (2006) of the static NS model. Therefore, the fact that no-dominance and no-arbitrage models affine models can hardly be distinguished empirically is consistent with a large body of evidence showing the empirical success of NS-based models. Conversely, the results in Fontaine, Feunou and Le (2014) provide a theoretical justification based on no-dominance of the empirical success of dynamic NS models.

Example 2 – Positive short rate It is easy to see that any choice of functionals with $m(\cdot) > 0$ guarantees that yields are positive. Fontaine, Feunou and Le (2014) advocate choosing specifications such that the one-period bond is non-decreasing on the $[0, 1]$ interval to guarantee positive forward rates. The logistic transformation is a natural choice: $B_t^{(1)} = 1/(1 + e^{s_t})$, implying that $r_t = log(1 + e^{s_t})$ is always positive. The logistic function captures the effect that the $\max$ function applied to the instantaneous rate in Black (1995) has on longer-maturity yields. Figure 14.11 shows the one period bond yields as a function of s_t. This is analogous to the difference between an option-like payoff, which exhibits a kink, and its price one-period ahead, which is smooth. Building on this intuition, Fontaine, Feunou and Le (2014) use a linear specification for $g(x)$ and a minor generalization to the logistic function for $m(x)$:

$$\begin{aligned} g(x) &= Kx \\ m(x) &= \theta_t \log(1 + \exp((\delta_0 + \delta'_1 x)/\theta_t), \end{aligned} \tag{14.74}$$

where θ_t is given in terms of the macro variables $\mathcal{M}_{t-1}$:

$$\theta_t = \theta_0 + \theta'_1 \mathcal{M}_{t-1}, \tag{14.75}$$

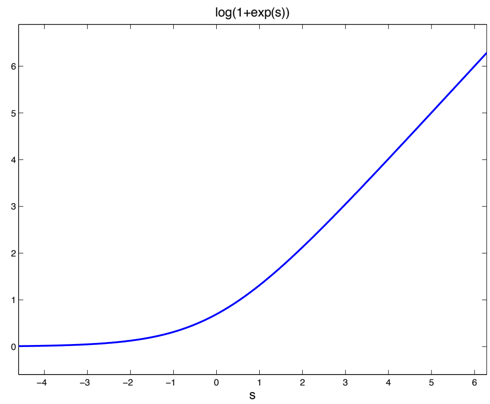


FIGURE 14.11 **Logistic transformation of the short rate.**

and controls how quickly the transformation affects the observed short rate r_t as the shadow rate s approaches its lower bound. The short rate in (14.74) can be inverted and, therefore, the shadow rate $s = \delta_0 + \delta_1' x$ can be inverted from a suitable number of forward rate portfolios: estimation can proceed very quickly without the need for a cumbersome filtering step. No-dominance term structure models allow for flexible time-series specifications. There is no restriction either on the choice of functions $m(\cdot)$ and $g(\cdot)$, except for $g(x) \in \underline{\mathrm{X}}$ (and technical conditions for the purpose of estimation).

14.4.4 RECENT EMPIRICAL RESULTS

Fontaine, Feunou and Le (2014) perform an early implementation of a truncated, or Black's, version of the benchmark Gaussian DTSM (B-DTSM), to Japanese data in their case. They find that a two-factor B-DTSM provides a good fit of the yield cross-section between 1996 and 2006. Kim and Singleton (2012) also study Japanese yield data and compare 2-factor DTSMs with the affine and quadratic short rate, and Black's version of these models, where the shadow rate is either affine or quadratic. Their results favor B-DTSMs. Empirical work has turned its attention to the United States, where the short rate reached its lower bound at the end of 2008. Most of this work focuses on variants of the B-DTSMs. While some of these results are still preliminary, the results suggest the following empirical regularities.

First, there is ample evidence that $\hat{s}_t$ has different properties—level, persistence, shocks—depending on the model specification. Hence, estimates of the shadow rate $\hat{s}_t$ may provide a misleading measure of the stance of monetary policy. There seems to be a consensus to use model implications that are more directly tied to observable variables. For instance, Bauer and Rudebusch (2013) advocate computing the *modal path* of the future policy rate and, in particular, calculating the time until the modal path crosses the lower bound as a measure of monetary stance.

Second, incorporating the lower bound is sufficient to capture (i) changes in yield volatilities across the term structure, (ii) changes in correlation between yields, and (iii) dampening of the risk premium when the short rate has reached its lower bound. This can be seen from estimation results for the following specification (see Fontaine, Feunou and Le (2014) for details). First, bond prices are given by Equations 14.69 and 14.74. Second, the risk factor $\mathcal{P}$ has autoregressive dynamics,

$$\mathcal{P}_{t+1} = K_0^{\mathbb{P}} + K_1^{\mathbb{P}}\mathcal{P}_t + \sqrt{\Sigma_t}\epsilon_{t+1}. \tag{14.76}$$

Third, risk factors' innovations have a dynamic conditional correlation (DCC) specification, as in Engle (2002), to capture the time-varying variances and correlations across yields. Formally, the covariance matrix is given by

$$\Sigma_t = D_t C_t D_t, \tag{14.77}$$

where $D_t = \text{diag}\left(\sigma_{1,t}, ..., \sigma_{K,t}\right)$ is a diagonal matrix containing the conditional volatility of each risk factor and C_t is the conditional correlation matrix. Each element of D_t follows an EGARCH$(1,1)$:

$$\log\sigma_{k,t}^2 = \omega_k + \beta_k \log\sigma_{k,t-1}^2 + \alpha_k Z_{k,t} + \kappa_k\left(|Z_{k,t}| - E[|Z_{k,t}|]\right), \tag{14.78}$$

where $Z_{k,t} = \epsilon_{k,t}/\sigma_{k,t}$ is the standardized innovation in the k-th risk factor. The conditional volatility exhibits a "leverage" effect whenever $\kappa \neq 0$: the conditional volatility is then correlated with the level of the risk factor. The correlation matrix C_t is defined via the autoregressive matrix Q_t:

$$Q_t = (1 - a - b)\,\bar{Q} + a\left(Z_t Z_t'\right) + bQ_{t-1}, \tag{14.79}$$

which is a symmetric positive definite matrix if $a > 0$, $b > 0$, $a + b < 1$ and $\bar{Q}$ is symmetric positive definite. The following standardization yields the correlation matrix:

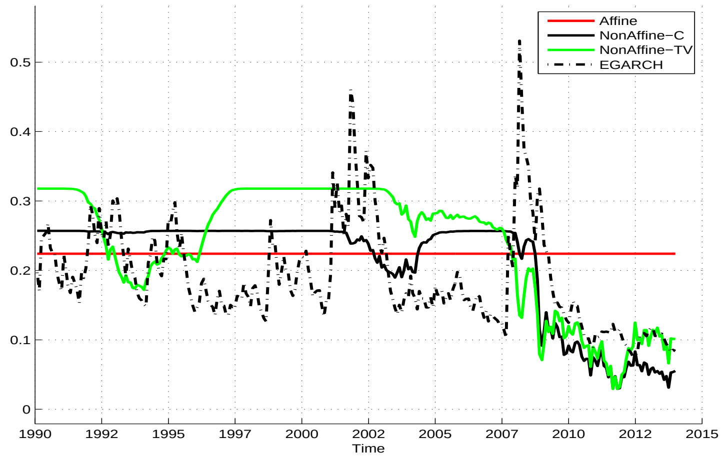
$$C_t = Q_t./\left(\text{diag}\left[Q_t\right]^{-1/2} \otimes \text{diag}\left[Q_t\right]^{-1/2}\right), \tag{14.80}$$

where $./$ is the element-by-element division of the matrix, $\otimes$ is the Kronecker product and $\text{diag}\left[Q_t\right]$ is the vector whose elements are the main diagonal of Q_t. Needless to say, the EGARCH, leverage and conditional correlations cannot be integrated within affine no-arbitrage term structure models.

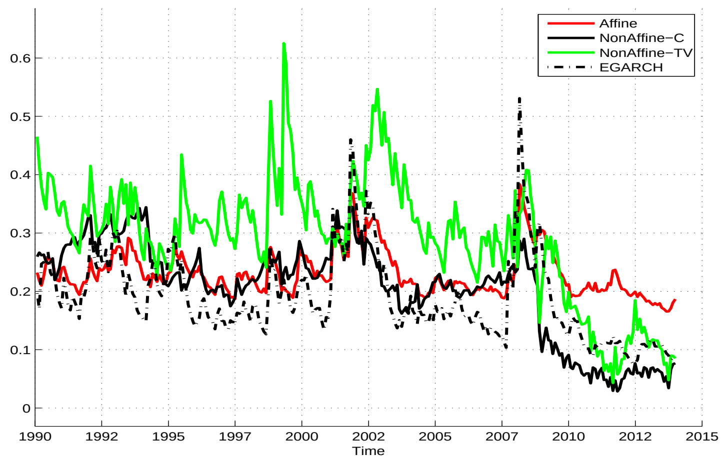
Panel (a) of Figure 14.12 compares the conditional volatility of the 2-year yield from a range no-dominance models. Panel (a) reports the conditional volatility for three models where the covariance is kept constant: one *affine* model, one *non-affine constant (C)* model where the θ parameter is constant and one *non-affine time-varying (TV)* model where the θ_t parameter is time-varying. The conditional volatility of the 5-year yield measured from the univariate EGARCH is higher in periods when the short rate is decreasing: toward

FIGURE 14.12 **Conditional Volatility of the 2-year Yield**

Panel (a) Constant covariance matrix Σ



Panel (b) Time-varying covariance matrix Σ_t



Conditional volatility of the 2-year yield from an unrestricted EGARCH and from no-dominance models without EGARCH or DCC in the risk factors innovations (Panel (a) and from no-dominance models with EGARCH and DCC (Panel (b)) . *Affine* is a no-dominance affine specification, *non-affine-C* is a specification enforcing the lower bound but with constant θ parameters, *non-affine-TV* is a specification enforcing the lower bound with time-varying θ_t parameters. US monthly data 1990–2014.

the end of the easing cycle in 1991, early in 1995, in 2001 and in 2007–2008. The volatility is constant and close to 22% annually in the *affine* model.

The *non-affine* models can only produce a lower volatility when yields are sufficiently close to the lower bounds. In particular, the model volatilities are very close to the EGARCH volatilities after 2008, when the short rate reaches its lower bound in the United States. By design, these models do not produce time-varying volatilities from the lower bounds and cannot match the EGARCH evidence.

Panel (b) of Figure 14.12 provides the same comparison for models with DCC. What is key for our discussion is that (i) the *affine* model with DCC cannot match the conditional volatility of the 2-year yield after 2008, and that (ii) adding more complicated dynamics does not change the volatilities in the *non-affine* model after 2008. We conclude that the compression due to the lower bound is the dominant feature of the recent sample and it is sufficient to capture the compressed volatility since 2009.

This result extends to the correlation between yields. We expect the correlation structure between yields to tighten due to the lower bound. Figure 14.13 compares the share of the conditional variance of all yields explained by the first principal component (where the principal component is computed using the full sample). Panel (a) reports results for models where the covariance of innovations is constant. The share of variance explained by the first component is constant in the *affine* model, close to 90%. Again, the *non-affine* models can only increase this share when yields are close enough to the lower bound. In particular, the share of variance explained remains close to 95% in the period following 2008.

Panel (b) reports results for models where the covariance of innovations is time-varying. The conclusions reached from Figure 14.12 remain: (i) the share is much lower after 2008 in the *affine* model even allowing for DCC, and (ii) allowing for DCC does not change the share of variance explained by the first component in the *non-affine* model. We conclude that the compression due to the lower bound is the dominant feature of the recent sample and it is sufficient to capture the increased correlation since 2009.

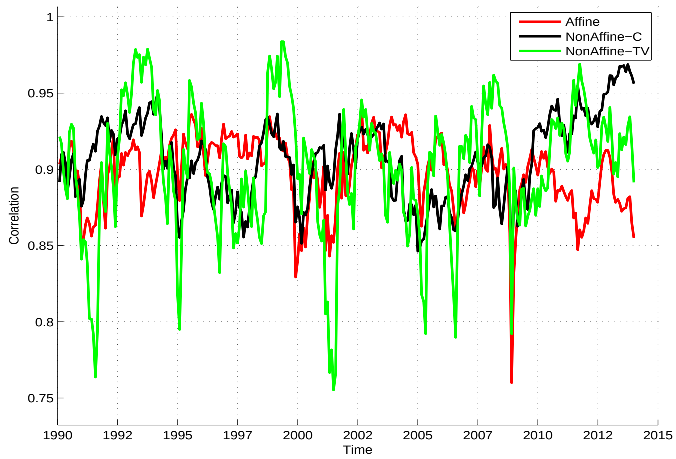
14.5 Conclusion

We have reviewed three old topics that once were pervasive in macro-finance fixed income research: liquidity, learning and the lower bound. The dispersion of bond prices, funding risk and short-selling pressures predict risk premia across fixed-income markets such as Treasury, LIBOR, swap, and agency and corporate bond markets. Since most empirical studies of term structure models were conducted with smoothed zero-coupon bond yields, these liquidity premia were obscured. Therefore, going back to coupon bond prices brings important information but entails some loss of model tractability.

FIGURE 14.13 Variance Share of the First Principal Component

Panel (a) Constant covariance matrix Σ

Panel (b) Time-varying covariance matrix Σ_t



Share of yields conditional variance explained by the first principal component extracted from the full sample from No-Dominance models without EGARCH or DCC in the risk factors' innovations (Panel (a)) and from No-Dominance models with EGARCH and DCC (Panel (b)). *Affine* is a no-dominance affine specification, *non-affine-C* is a specification enforcing the lower bound but with constant θ parameters, *non-affine-TV* is a specification enforcing the lower bound with time-varying θ_t parameters. US monthly data 1990–2014.

Information from surveys on forecasts by economic agents and professionals plays an important role in explaining expected returns. Expectational errors explain part of the statistical bond premia and may be interpreted as waves of optimism or pessimism that make agents depart from their rational expectations. The role played by expectation shocks has not been fully explored and is promising area for future research. Both the magnitude and the dynamics of bond premia otherwise obtained in no-arbitrage or equilibrium models with rational expectations will likely be affected in a substantial way.

Estimation of term structure models over the past fifty years or so faces with a serious empirical challenge, with very high rates in the beginning of the eighties and rates close to zero in the recent period. The evidence so far suggests that any model imposing a lower bound on yields can capture the compression in volatility and the increased correlations when yields are close to the bound. On the other, matching the bond risk premium and volatility when yields are away from their bound remains a challenge for most implementation. Progress has been made, but the work has just begun. In sum, the return to the forefront of old topics in fixed-income research has produced a stimulating agenda, which holds promise for continued rapid advancement in the future.

Acknowledgements

We thank Bruno Feunou for several helpful discussions, we thank the Pietro Veronesi, Florian Pelgrin, Jean-Paul Rennes, Francisco Rivadeneyra Sanchez for comments on earlier drafts of this chapter, and we thank Glen Gleeleyside for editorial assistance.

REFERENCES

Adrian, T. and Shin H. (2008), *Liquidity and Leverage*, Staff Report 361, Federal Reserve Bank of New York.

Amihud, Y. and Mendelson, H. (1991), *Liquidity, Maturity and the Yields on U.S. Treasury Securities*, Journal of Finance, vol. 46(4), pp. 1411Ű1425.

Bacchetta, P., Mertens, E. and van Wincoop, E. (2009), *Predictability in Financial Markets: What do Survey Expectations Tell Us?*, Journal of International Money and Finance, vol. 28(3), pp. 406-426.

Banerjee, S. and Graveline, J. (2013), *The Cost of Short-Selling Liquid Securities*, Journal of Finance, vol. 68(2), pp. 637-664.

Bansal, R. and Yaron, A. (2004), *Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles*, Journal of Finance vol. 59(4), pp. 1481-1509.

Bartolini, L., Hilton, S., Sundaresan, S. and Tonetti, C. (2011), *Collateral Values by Asset Class: Evidence from Primary Security Dealers*, Review of Financial Studies, vol. 24(1), pp.248-278.

Bauer, M. D. and Rudebusch, G. D. (2013), *Monetary Policy Expectations at the Zero Lower Bound*, Working paper, Federal Reserve Bank of San Francisco.

Black, F. (1995), *Interest Rates as Options*, Journal of Finance, vol. 50(5), pp. 1371-1376.

Bliss, R. (1997), *Testing Term Structure Estimation Methods*, Advances in Futures and Options Research, vol. 9, pp. 197-231.

Bulkley, G. and Giordani, P. (2011), *Structural Breaks, Parameter Uncertainty, and Term Structure Puzzles*, Journal of Financial Economics, vol. 102(1), pp. 222-232.

Buraschi, A., and Menini, D. (2002), *Liquidity Risk and Specialness*, Journal of Financial Economics, vol. 64(2), pp. 243-284.

Campbell, J. and Shiller, R. (1991), *Yield Spreads and Interest Rate Movements: A BirdŠs Eye View*, Review of Economic Studies, vol. 58, pp. 495Ű514.

Chen, X., Sun, Z., Yao, T. and Yu, T. (2014), *In Search of Habitat*, Working paper.

Cheria, J., Jacquier, E., and Jarrow, R. (2004), *A Model of the Convenience Yields in On-the-Run Treasuries*, Review of Derivatives Research, vol. 7, pp. 79-97.

Chernov, M. and Mueller, P. (2011), *The Term Structure of Inflation Expectations*, Journal of Financial Economics, vol. 106(2), pp. 367-394.

Christensen, J., Diebold, F. and Rudebusch, G. (2011), *The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models*, Journal of Econometrics, vol. 164(1), pp. 4-20.

Chun, A. L. (2011), *Expectations, Bond Yields, and Monetary Policy*, Review of Financial Studies, vol. 24(1), pp. 208-247.

Cieslak, A. and P. Povala (2013), *Expecting the Fed*, Working paper, Northwestern University.

Cochrane, J. and Piazzesi, M. (2005), *Bond Risk Premia*, American Economic Review, vol. 95(1), pp. 138-160.

Collin-Dufresne, P., Johannes, M. and Lochstoer, L. (2013), *Parameter Learning in General Equilibrium: The Asset Pricing Implications*, Working paper, Columbia Business School.

Cornell, B. and Shapiro, A. C. (1989), *The Mispricing of US Treasury Bonds: A Case Study*, Review of Financial Studies, vol. 2(3), pp. 297-310.

Coibion, O. and Gorodnichenko, Y. (2012), *What can Survey Forecasts tell us about Information Rigidities?* Journal of Political Economy, vol. 120(1), pp. 116-159.

Cox, J., Ingersoll, J. and Ross, S. (1985), *An Intertemporal General Equilibrium Model of Asset Prices*, Econometrica, vol. 53(2), pp. 363-384.

Culbertson, J. M. (1957), *The Term structure of Interest Rates*, Quarterly Journal of Economics, vol. 71(4), pp. 485-517.

Dai, Q. and Singleton, K. (2000), *Specification Analysis of Affine Term Structure Models*, Journal of Finance, vol. 55(5), pp. 1943-1978.

Dai, Q. and Singleton, K. (2002), *Expectation Puzzles, Time-Varying Risk premia and Affine Model of the Term Structure*, Journal of Financial Economics, vol. 63(3), pp. 415-441.

Dai, Q. and Singleton, K. (2003), *Fixed Income Pricing*, Chapter 20, pp. 1207-1246, in Handbook of Economics and Finance, North Holland.

Del Negro, M. and Eusepi, S. (2011), *Fitting Observed Inflation Expectations*, Journal of Economic Dynamics and Control, vol. 35(12), pp. 2105-2131.

Detemple, J. B. (1986), *Asset Pricing in a Production Economy with Incomplete Information*, Journal of Finance, vol. 41(2), pp. 383-391.

Diebold, F. and Li, C. (2006), *Forecasting the Term Structure of Government Bond Yields*, Journal of Economic Dynamics and Control, vol. 130(2), pp. 337-364.

Dothan, M. U. and Feldman, D. (1986), *Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy*, Journal of Finance, vol. 41(2), pp. 369-382.

Duarte, J. (2008), *The Causal Effect of Mortgage Refinancing on Interest Rate Volatility: Empirical Evidence and Theoretical Implications*, Review of Financial Studies, vol. 21(4), pp. 1689-1731.

Duffee, G. (1996), *Idiosyncratic Variation of Treasury Bill Yields*, Journal of Finance, vol. 51(2), pp. 527-551.

Duffee, G. (2002), *Term Premia and Interest Rate Forecasts in Affine Model*, Journal of Finance, vol. 57, pp. 405-443.

Duffee, G. (2013), *Bond Pricing and the Macroeconomy*, Chapter 13, pp. 907-967, in Handbook of Economics and Finance, Elsevier.

Duffie, D. (1996), *Special Repo Rates*, Journal of Finance, vol. 51(2), pp. 493-526.

Duffie, D. and Kan, R. (1996), *A Yield-Factor Model of Interest Rates*, Mathematical Finance, vol. 6(4), pp. 379-406.

Engle, R. (2002), *Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models*, Journal of Business and Economic Statistics, vol. 17(3), pp. 339Ű350.

Fama, E. and Bliss, R. (1987), *The Information in Long-Maturity Forward Rates*, American Economic Review, vol. 77(4), pp. 680-692.

Fernald, J. D., Keane, F. and Mosser, P.C. (1994), *Mortgage Security Hedging and the Yield Curve*, Quarterly Review - Federal Reserve Bank of New York, vol. 19, pp. 92-100.

Feunou, B. and Fontaine, J.-S. (2014), *Non-Markov Gaussian Term Structure Models: The Case of Inflation*, Review of Finance vol. 18(5), pp. 1953-2001.

Feunou, B., Fontaine, J.-S., and Le, A. (2014), *Term Structure Modeling when Monetary Policy is Unconventional*, Working paper, Bank of Canada.

Fontaine, J.-S. and Garcia, R. (2012), *Bond Liquidity Premia*, Review of Financial Studies, vol. 25(4), pp. 1207-1254.

Fontaine, J.-S., Garcia, R. and Gungor, S. (2015), *Funding Liquidity, Market Liquidity, and the Cross-section of Stock Returns*, Working paper, Bank of Canada.

Friedman, B. M. (1980), *Survey Evidence on the Rationality of Interest Rate Expectations*, Journal of Monetary Economics, vol. 6(4), 453-465.

Froot, K. A. (1989), *New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates*, The Journal of Finance, vol. 44(2), pp. 283-305.

Gabaix, X., Krishnamurthy, A. and Vigneron, O. (2007), *Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market*. Journal of Finance, vol. 62(2), pp. 557-595.

Garbade, K. (1984), *Analyzing the Structure of Treasury Yields: Duration, Coupon and Liquidity Effects*, Topics in Money and Securities Markets, Bankers Trust Co.

Garbade, K. and Silber, W. L. (1976), *Price Dispersion in the Government Securities Market*, Journal of Political Economy, vol. 84(4), pp. 721-740.

Garcia, R. and Luger, R. (2012), *Risk Aversion, Intertemporal Substitution, and the Term Structure of Interest Rates*, Journal of Applied Econometrics, vol. 27(6), pp. 1013-1036.

Gourieroux, C. and Jasiak, J. (2006), *Autoregressive Gamma Process*, Journal of Forecasting, vol. 25(2), pp 129-152.

Gourinchas, P.-O.and Tornell, A. (2004), *Exchange Rate Puzzles and Distorted Beliefs*, Journal of International Economics, vol. 64(2), pp. 303-333.

Graveline, J. and McBrady, M.R. (2011), *Who Makes On-the-Run Treasuries Special?*, Journal of Financial Intermediation, vol. 20(4), pp. 620-632.

Greenwood, R. and Vayanos, D. (2014), *Bond Supply and Excess Bond Returns*, Review of Financial Studies, vol. 27(3), pp. 663-713.

Guibaud, S., Nosbusch, Y. and Vayanos D. (2013), *Bond Market Clienteles, the Yield Curve, and the Optimal Maturity Structure of Government Debt*, Review of Financial, vol. Studies 26(8), pp. 1914Ű1961.

Gurkaynak, R., Sack, B. and Wright, J. (2006), *The U.S. Treasury Curve: 1961 to the Present*, Working paper 2006-28, Federal Reserve Board.

Gürkaynak, R. and Wright, J. H. (2012), *Macroeconomics and the Term Structure*, Journal of Economic Literature, vol. 50(2), pp. 331-367.

Hanson, S. (2014), *Mortgage Convexity*, Journal of Financial Economics, vol. 113(2), pp. 270-299.

Ichiue, H. and Ueno, Y. (2007), *Equilibrium Interest Rates and the Yield Curve in a Low Interest Rate Environment*, Working Paper 2007-E-18, Bank of Japan.

Jegadeesh, N. (1993), *Treasury Auction Bids and the Salomon Squeeze*, Journal of Finance, vol. 48(4), pp. 1403-1419.

Jongen, R., Verschoor, W. and Wolff, C. (2005), *Time Variation in Term Premia: International Evidence*, Discussion paper, CEPR.

Jordan, B. and Jordan, S. (1997), *Special Repo Rates: An Empirical Analysis*, Journal of Finance, vol. 52(5), 2051-2072.

Joslin, S. and Le, A. (2011), *Interest Rate Volatility and No-Arbitrage Affine Term Structure Models*, Working paper, Kenan-Flagler Business School.

Joslin, S., Singleton, K. J., and Zhu, H. (2011), *A New Perspective on Gaussian Dynamic Term Structure Models*, Review of Financial Studies, vol. 24(3), pp. 926Ű970.

Joslin, S., Priebsch, M. and Singleton, K. J. (2014), *Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks*, Journal of Finance, vol. 69(3), pp. 1197-1233.

Kamara, A. (1994), *Liquidity, Taxes, and Short-term Treasury Yields*, Journal of Financial and Quantitative Analysis, vol. 29(3), pp. 403-417.

Kim, D. H. and Orphanides, A. (2012), *Term Structure Estimation with Survey Data on Interest Rate Forecasts*, Journal of Financial and Quantitative Analysis, vol. 47(1), pp. 241-272.

Kim, D. H. and Singleton, K. J. (2012), *Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields*, Journal of Econometrics, vol. 170(1), pp 32-49.

Krippner, L. (2012), *Modifying Gaussian Term Structure Models when Interest Rates are Near the Zero Lower Bound*, Working paper, Reserve Bank of New Zealand.

Krippner, L. (2013), *A Theoretical Foundation for the Nelson-Siegel Class of Yield Curve Models*, Journal of Applied Econometrics, vol. 30(1), pp. 97-118.

Krishnamurthy, A. (2002), *The Bond/Old-Bond Spread*, Journal of Financial Economics, vol. 66(2-3), pp. 463-506.

Le, A., Singleton, K.J. and Dai Q. (2010), *Discrete-Time Affine $\mathbb{Q}$-Term Structure Models with Generalized Market Prices of Risk*, Review of Financial Studies, vol. 23(5), pp. 2184-2227.

Liu, J. and Longstaff, F. A. (2004), *Losing Money on Arbitrage: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities*, Review of Financial Studies, vol. 17(3), pp. 611Ű641.

Longstaff, F. (2004), *The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices*, Journal of Business, vol. 77(3), pp. 511Ű526.

Lou, D., Yan, H., and Zhang, J. (2013), *Anticipated and Repeated Shocks in Liquid Markets*, Review of Financial Studies, vol. 26(8), pp. 1891-1912.

Malkhozov, A., Mueller, P., Vedolin, A. and Venter, P. (2014), *Mortgage Risk and the Yield Curve*, Working paper, McGill University.

Markiewicz, A. and Pick, A. (2014), *Adaptive Learning and Survey Data*, Journal of Economic Behavior & Organization, vol. 107(B), pp. 685-707.

Meiselman, D. (1962), *The Term Structure of Interest Rates*. Ford Foundation Dissertation Series, Prentice-Hall.

Milani, F. (2007), *Expectations, Learning and Macroeconomic Persistence*, Journal of Monetary Economics, vol. 54(7), pp. 2065-2082.

Modigliani, F. and Sutch R. (1967), *Debt Management and the Term Structure of Interest Rates: an Empirical Analysis of Recent Experience*, The Journal of Political Economy, vol. 75(4), 569-589.

Monfort, A., Pegoraro, F., Renne, J-P and Roussellet, G. (2014), *Staying at Zero with Affine Processes: A New Dynamic Term Structure Model*, Working paper, Banque de France.

Obstfeld, M. (1990), *Intertemporal Dependence, Impatience, and Dynamics*, Journal of Monetary Economics, vol. 26(1), pp. 45-75.

Ormeno, A. (2009), *Inflation Expectations, Learning, and DSGE Model Estimation*, Working paper, Universitat Pompeu Fabra.

Perli, R. and Sack, B. (2003), *Does Mortgage Hedging Amplify Movements in Long-Term Interest Rates?*, Journal of Fixed Income, vol. 13(3), pp. 7Ű17.

Piazzesi, M. (2005), *Affine Term Structure Models*, Chapter 12, pp. 691-766, in Handbook of Financial Econometrics, Elsevier.

Piazzesi, M., Salomao, J. and Schneider, M. (2013), *Trend and Cycle in Bond Premia*, Working paper, Stanford University.

Priebsch, M. (2013), *Computing Arbitrage-Free Yields in Multi-Factor Gaussian Shadow-Rate Term Structure Models*, Working paper, Federal Reserve Board.

Rosenbaum, S. (1961), *Moments of a Truncated Bivariate Normal Distribution*, Journal of the Royal Statistical Society, Series B (Methodological), vol. 23(2), pp. 405-408.

Sarig, O. and Warga, A. (1989), *Some Empirical Estimates of the Risk Structure of Interest Rates*, Journal of Finance, vol. 44(5), 1351-1360.

Shiller, R. and McCulloch, J. (1987), *The Term structure of Interest Rates*, Chapter 13, pp. 627-722, in Handbook of Monetary Economics, vol. 1, Elsevier.

Slobodyan, S. and Wouters, R. (2012), *Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models*, American Economic Journal: Macroeconomics, vol. 4(2), pp. 65-101.

Tobin, J. (1958), *Estimation of Relationships for Limited Dependent Variables*, Econometrica, vol. 26(1), pp. 24-36.

Vayanos, D. and Vila, J.-L. (2009), *A Preferred-Habitat Model of the Term Structure of Interest Rates*, Working paper, W15487, NBER.

Vayanos, D. and Weill, P.-O. (2006), *A Search-Based Theory of the On-the-Run Phenomenon*, Journal of Finance, vol. 63(3), pp. 1361-1398.

Warga, A. (1992), *Bond Returns, Liquidity and Missing Data*, Journal of Financial and Quantitative Analysis, vol. 27, 605-617.

Wu, C. and Xia, F. D. (2013), *Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound*, Working paper, University of Chicago Booth School of Business.

Moments of Truncated Bivariate Distribution

As in the text, define

$$\begin{aligned}\mu_{i|t} &= E_t^{\mathbb{Q}}[s_{t+i}] \\ \sigma_{i|t}^2 &= Var_t^{\mathbb{Q}}[s_{t+i}] \\ \sigma_{i,n|t} &= Cov_t^{\mathbb{Q}}(s_{t+i}, s_{t+n}),\end{aligned} \tag{14.81}$$

so that

$$\begin{pmatrix} s_{t+i} \\ s_{t+n} \end{pmatrix} |X_t \sim N^{\mathbb{Q}} \left[\begin{pmatrix} \mu_{i|t} \\ \mu_{n|t} \end{pmatrix}, \quad \begin{pmatrix} \sigma_{i|t}^2 & \sigma_{i,n|t} \\ \sigma_{i,n|t} & \sigma_{n|t}^2 \end{pmatrix} \right] \quad 0 \leq i \leq n, \tag{14.82}$$

for $n = 1, 2, \ldots$. The first two moments of the truncated distribution are given by

$$E_t[\max(0, s_{t+i})] = \mu_{i|t}\Phi(\hat{\mu}_{i|t}) + \sigma_{i|t}\phi(\hat{\mu}_{i|t}), \tag{14.83}$$

and

$$\begin{aligned} E_t[\,\max(0, s_{t+i}) \max(0, s_{t+n})] &= (\mu_{i|t}\mu_{n|t} + \sigma_{i,n|t})\Phi_2^d\left(\hat{\mu}_{i|t}, \hat{\mu}_{n|t}; \rho_{i,n|t}\right) \\ &+ \sigma_{n|t}\mu_{i|t}\phi(\hat{\mu}_{n|t})\Phi\left(\frac{\hat{\mu}_{n|t} - \hat{\mu}_{t+i|t}\rho_{i,n|t}}{\sqrt{1-\rho_{i,n|t}^2}}\right) \\ &+ \sigma_{i|t}\mu_{n|t}\phi(\hat{\mu}_{i|t})\Phi\left(\frac{\hat{\mu}_{i|t} - \hat{\mu}_{t+n|t}\rho_{i,n|t}}{\sqrt{1-\rho_{i,n|t}^2}}\right) \\ &+ \sigma_{1|t}\sigma_{n|t}\sqrt{\frac{1-\rho_{i,n|t}^2}{2\pi}}\phi\left(\sqrt{\frac{\hat{\mu}_{i|t}^2 - 2\rho_{i,n|t}\hat{\mu}_{i|t}\hat{\mu}_{n|t} + \hat{\mu}_{n|t}^2}{1-\rho_{i,n|t}^2}}\right), \end{aligned} \tag{14.84}$$

where we defined $\hat{\mu}_{i|t} \equiv \mu_{i|t}/\sigma_{i,n|t}$ and $\rho_{i,n|t} \equiv \frac{\sigma_{i,n|t}}{\sigma_{i|t}\sigma_{n|t}}$; where $\Phi(\cdot)$ is the normal cumulative distribution function, $\phi(\cdot)$ is the normal probability distribution function; and where

$$\Phi_2^d(z_1, z_2; \rho) = 1 - \Phi(z_1) - \Phi(z_2) + \Phi_2(z_1, z_2; \rho), \tag{14.85}$$

with $\Phi_2(z_1, z_2; \rho)$ the bivariate normal cumulative distribution function. The proof follows directly from Lemmas A.1 and A.2 in Appendix A.2 of Priebsch (2013).