Review of Derivatives Research (2020) 23:191–226
https://doi.org/10.1007/s11147-019-09162-z

Yield curves from different bond data sets

Antonio Díaz[1] · Francisco Jareño[1] · Eliseo Navarro[2]

Published online: 5 July 2019

Abstract
It is well known that zero coupon rates are not observable variables. Their estimation process may be cumbersome and time consuming. We explore the extent to which the set of security prices used in the yield curve construction of three popular interest rate datasets (from the Federal Reserve Board, the US Department of the Treasury, and Bloomberg) may determine the results of different analyses. Using the same US Treasury prices from GovPX and applying the same fitting technique, we estimate zero coupon rates using different baskets of assets, i.e., including/excluding bills, on-the-run, and off-the-run bonds, attempting to mimic those used by each data providers. To illustrate the uncertainty surrounding these alternatives representations of the underlying yield curve, we examine common uses of these data sets in pricing, risk management and macroeconomic purposes. We find significant and sometime overwhelming differences in the volatility term structure, the pricing of interest rate derivatives, and the correlations among different forward rates particularly in both ends of the yield curve. Relevant implications are also observed on a classic test of the expectations hypothesis. The simplest asset basket, which only includes the on-the-run bills and bonds, is probably the one with the best results.

Keywords Term structure of interest rates · Yield curve datasets · Volatility term structure · Forward rates · Expectations hypothesis

JEL Classification E43 · F31 · G12 · G13 · G15

✉ Antonio Díaz
Antonio.Diaz@uclm.es

Francisco Jareño
Francisco.Jareno@uclm.es

Eliseo Navarro
Eliseo.Navarro@uah.es

1 Departamento de Análisis Económico y Finanzas, Facultad C. Económicas y Empresariales, Universidad de Castilla-La Mancha, Plaza de la Universidad, 1, 02071 Albacete, Spain

2 Departamento de Economía y Dirección de Empresas, Edificio C. Económicas, Universidad de Alcalá, Empr. y Tur., Pz. de la Victoria, s/n, 28802 Alcalá de Henares, Spain

1 Introduction

It is unnecessary to justify the importance of the key role that risk-free zero-coupon yields play in all fields of finance for both practitioners and researchers. However, on the whole, these risk-free yields are not directly observable. Instead, they must be estimated from risk-free bond markets. As is well known, the estimation of these rates can sometimes be cumbersome and time consuming. Thus, when these data are needed, professionals or researchers usually download them from database providers instead of performing their own estimations. This is the reason that some data sets—-particularly those that are either publicly available or offered by primary financial data providers—have become quite popular. In our study, we examine three of the most used data sets: the zero-coupon yield curves provided by the Federal Reserve Board (FRB), by the US Department of Treasury (DoT), and by Bloomberg.

There are two primary sources of the discrepancies among yield curve data sets: (1) the model and the numerical techniques employed to estimate zero coupon rates; and (2) the basket of securities used as input for the estimation. Users of yield curves (traders, portfolio managers, researchers, policymakers, etc.) are all aware of the fact that constructed zero-coupon yields are "just" representations of the underlying yield curve. Most final users of these yield curves chose one among the different available data sets in function of the trade-off between smoothness and goodness of fit, i.e., the decision about the yield curve to use as input is based on the first primary source of discrepancies. Final users interested in purposes of macroeconomic and monetary policy analysis may consider that minor discrepancies across different yield databases are entirely irrelevant for the questions at hand. In contrast, they demand yield curves that are rigid and stable over time. For other final users, such as traders and portfolio managers tasked with pricing contracts that are related to, but different from, those currently observed in the market, the error or uncertainty surrounding the zero-coupon yield estimates may seriously complicate inference. In these cases, the accuracy is paramount. However, it is not clear if all researchers and practitioners do realize the real implications resulting from using one of the alternative yield curve data sets according with the concrete set of Treasury securities considered in the fitting process by the yield curve provider. Concerns about the appropriateness of the criterion to include/exclude assets applied by the provider may be beyond the scope of most final users of these yield curves.

Considerable research effort has been devoted to finding an accurate and consistent technique for estimating the term structure of interest rates (e.g., Bliss 1996; Bolder and Gusba 2002; Jordan and Mansi 2003; Bolder et al. 2004; Yallup 2012). These works primarily focus on the ability of alternative models and methodologies to replicate sets of bond prices; however, they do not pay attention to other variables derived from the analysis of the time series of data obtained through these alternative methods. The recent working paper by Andreasen et al. (2017) compare the properties of the risk-neutral parameters obtained from arbitrage-free dynamic term structure models estimated using a two-step approach, i.e., interpolating from constant-maturity zero-coupon yields, and a one-step approach. In addition, they observe nonnegligible measurement errors in yields from using two curve-fitting techniques from the same data set of Canadian government bond prices.

To the best of our knowledge, the potential implications of the second issue, i.e., the set of securities used in the yield curve construction, have not yet been explicitly studied. It should be noted that different financial information providers use different prices (transaction prices, quoted prices or market yields to maturity) and different sets of risk-free debt instruments (including/excluding bills, on-the-run bonds,[1] off-the-run bonds, old callable bonds, etc.) to estimate zero-coupon bond yields. Previous literature observes that different liquidity and specialness characteristics affect bond pricing. In fact, the on/off-the-run phenomenon is well known and has been studied by many authors. It should be expected that yield curves obtained by different providers, and therefore their statistics, are different. However, as far as we know, this is the first study to isolate the effects and to evaluate consequences of the decisions taken by yield curve providers about the basket of considered assets.

The aim of this paper is to explore the extent to which the set of bond prices used in the yield curve construction of three popular interest rate datasets (the FRB, the DoT, and Bloomberg) may determine the results of different analyses. To illustrate the uncertainty surrounding these alternatives representations of the underlying yield curve, we examine common uses of these data sets in pricing, risk management and macroeconomic purposes.

In our study, we first examine daily estimates of the three popular zero-coupon yield curves in the period 1996–2006.[2] The providers of these three data sets use different models and techniques, different market data, and different baskets of assets. Therefore, the direct comparison among these data sets does not make it possible to isolate the effect of each source of discrepancy. To address this problem we proceed as follows: using intraday US Treasury quotes and trades from the same database (GovPX) and applying the same model (Svensson), we estimate zero coupon rates using different baskets of assets, attempting to mimic those used by each of the three data providers. The only difference among our estimates is the Treasuries included as input to estimate the zero-coupon yields. Instead of focusing on the ability of risk-free zero-coupon rates to replicate bond prices or yields, we examine the potential impact of the time series of zero-coupon rate estimates on variables such as interest rate volatilities or correlations among forward interest rates with different maturities. These variables have a key role in many financial issues, such as the implementation of interest rate models, e.g., Black or Heath–Jarrow–Morton models, in the pricing of interest rate and volatility derivatives, in many product valuations where correlations among forward rates are crucial, e.g., swaptions, in analyses of market expectations about future short-term interest rates, and in risk management purposes, e.g., Value-at-Risk and Expected Shortfall calculations.

From our four samples of zero-coupon interest rates obtained from different baskets of assets, we study potential implications on the estimates of other variables that imply using the time series of interest rates: the volatility term structure (VTS) and some

[1] The most recently issued security of a particular maturity.

[2] As commented below, we replicate the baskets of securities these providers use to fit the Yield Curve (YC). We get intraday market prices of the US Treasury securities from the GovPX dataset. Thus, our sample period is limited for the availability of reliable prices from GovPX. The quality of this dataset deteriorates progressively, and we are forced to discontinue the sampling period at the end of 2006. See more details in the "Appendix".

correlations between pairs of forward rates. To estimate interest rate volatilities, we use two alternative methods: simple standard deviations using a 3-month rolling window and the EGARCH(1,1) model. We observe statistically significant differences between samples in the behavior of volatilities and correlations. We illustrate the implications of these findings through two exercises. First, we calibrate the Black et al. (1990) model, widely used by the industry, to price callable bonds. Relevant discrepancies are obtained both in relative and absolute terms in the pricing of the embedded call options. Second, we replicate one of the most classical tests of the expectations hypothesis of the term structure of interest rates. The outcomes are partially dependent on the data employed.

Our main findings are that the universe of bond prices used in the construction of popular zero-coupon yields may have serious implications in common uses of these data sets. In addition, we show how the liquidity premium and the short-term interest rates complicate the effort to extract unique zero-coupon yields. There is no single "best" yields data set suitable to use for every purpose, but it should be used one or another yield curve data set depending on the specific application. Anyway, considering our findings and the lower computational cost when estimating the yield curve from only eleven securities, we can conclude that Sample D is probably the best option for many of the proposed uses.

The rest of this paper is organized as follows: Sect. 2 analyzes the primary characteristics of the external data sets that we examine. Section 3 describes the original bond prices dataset that we use, the fitting process for our own estimates using the Svensson method, and the different sets of assets to include. In Sect. 4, we perform an empirical analysis that consists of a study of the impact of the data choice on the spot rates, the term structure of volatilities and the pricing of callable bonds, the forward rate correlation, and a simple expectations hypothesis test. The last section summarizes results and conclusions.

2 Alternative yield curve datasets

In this paper, we consider three of the most popular datasets containing estimations of risk-free zero-coupon bond yields corresponding to the US Treasury market. Two of those datasets can be downloaded from the websites of the Federal Reserve Board (commented by Gürkaynak et al. 2007)[3] and the US Department of the Treasury ("Daily Treasury Yield Curve Rates"), (FRB and DoT datasets, henceforth, respectively). The latter set of interest rates also appear as "U.S. government securities. Treasury constant maturities" included in the Interest Rates H.15 series posted on the Fed's website.[4] In addition, financial data vendors such as Bloomberg and Thomson

[3] A daily updated spreadsheet can be downloaded from the FRB's website in the section "Finance and Economics Discussion Series" below the Gürkaynak et al. (2007)'s working paper (https://www.federalreserve.gov/econresdata/feds/2006/index.htm). For convenience, this dataset is referred to "the FRB dataset"; however, the website indicates as follows: "Note: This is not an official Federal Reserve statistical release.".

[4] Thus, this dataset appears simultaneously in the DoT's website (https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yield) and the Fed's website (https://www.federalreserve.gov/releases/h15/data.htm).

Reuters provide daily risk-free yield curves from government securities. We consider the Bloomberg zero-coupon yield curves (code F082).

The main features of these three data sets are described in Table 1, which focuses on the fitting technique, the market data used as inputs for the estimations and the basket of assets included in the estimations of the yield curves.

The first data set (FRB) uses a weighted version of Svensson's (1994) method. Gürkaynak et al. (2007) comment that they use end-of-the-day prices but do not identify the specific prices employed, that is, they do not indicate whether the prices are quoted prices (mid bid-ask, bid or ask prices) or trading prices (last trading price, average daily prices) and so forth. They exclude from their estimation all Treasury bills and the on-the-run and "first-off-the-run" issues of notes and bonds, bonds with less than 3 months to maturity, bonds with embedded options, 20-year bonds since 1996 and "other issues that [they] judgmentally exclude on an ad hoc basis." Most daily estimates incorporate more than one hundred observations.

The second data set studied in this paper (DoT) uses a quasi-cubic Hermite spline function[5] that passes exactly through the yields of a given set of securities that are calculated from composites of quotations obtained by the Federal Reserve Bank of New York. Therefore, DoT does not estimate a zero-coupon yield term structures but instead uses simple yield curves relating yields to maturity and terms to maturity. With respect to the basket of assets, DoT considers only on-the-run securities, including four maturities of the most recent auction bills (4-, 13-, 26- and 52-week), six maturities of just-issued notes and bonds (2-,3-, 5-. 7-, 10- and 30-year) and a composite rate in the 20-year maturity range. Note that as they consider on-the-run bills and bonds, the resulting yield curve can be interpreted as a par yield curve because these just-issued assets are traded near par. However, coupon bias and forward rate bias may appear and with them, differences between zero-coupon interest rates and par yields, especially for long maturities. It should be noted that daily yield curve estimations are obtained from eleven daily observations.

Finally, the third data set (F082) uses a piecewise linear function, although no more details about the concrete function are reported by Bloomberg. This data set uses the generic prices of all outstanding Treasury bonds to generate "Bloomberg fair value" curves for pricing most bonds that either are traded over-the-counter or are illiquid bonds.[6] The basket of assets in F082 is even larger than the one used in the FRB dataset, indicating that the number of daily observations is larger than two hundred.

We can see that these three data sets (each one with virtues and flaws) differ considerably in the model and methodology employed to estimate zero-coupon bond yields. In addition, the set and number of assets used to estimate the zero-coupon yields vary considerably, ranging from eleven to more than two hundred securities. Moreover, the differential characteristics of the assets employed may have a considerable impact on both the level and the shape of the estimated yield curves.

[5] No more details are reported about the specific functions used to estimate the yield curve.

[6] Bloomberg terminal mentions: "The yield curve is built daily with bonds that have either Bloomberg Generic (BGN) prices, supplemental proprietary contributor prices or both. The bonds are subject to option-adjusted spread (OAS) analysis and the curve is adjusted to generate a best fit. (…) Bloomberg Generic Price (BGN) is Bloomberg's market consensus price (…) Bloomberg Generic Prices are calculated by using prices contributed to Bloomberg and any other information that we consider relevant"

Table 1 Primary characteristics of the considered yield curve datasets

	Fitting technique	Market data	Bills	On-the-run bonds and notes	Remaining straight bonds and notes	Non-straight bonds (callable…)	Shortest maturity
Gürkaynak et al. (2007) in Federal Reserve Board (FRB)	Weighted Svensson (1994)	End-of-day prices	No	No	Yes (excluding the first off-the-run)	No	3-month
U.S. Department of the Treasury (DoT)	Quasi-cubic hermite spline function	Close of business bid yields-to-maturity	Yes (only 4-, 13-, 26-, 52-weeks)	Yes (only 2-, 3-, 5-, 7-, 10-, *20*-, 30-years)	No	No	1-month
Bloomberg (F082)	Piecewise linear function	Bloomberg generic prices (quotes over a time window), supplemental proprietary contributor prices or both	No	Yes	Yes	Yes	Not reported

It is well known that the securities traded in the Treasury market have some important features with significant effects on yields, including optionality, coupon size with its corresponding tax effects, remaining maturity and liquidity. In the case of the optionality, Bloomberg includes old outstanding 30-year callable bonds in the yield curve estimations. For most purposes, the inclusion of these bonds in the estimation can cause unnecessary noise, breaking the sample's homogeneity. The attention paid to each maturity tranche, the number of securities included in the fitting process and the relative number of bonds with longest maturities are relevant aspects in the fitting process. In this sense, the relevance assigned by the yield-curve data sets to maturities shorter than 1 year or longer than 10 years is markedly different. Four of the eleven maturities considered by the DoT fall under the first range, whereas three of the maturities are 10 years or longer. In contrast, the FRB excludes all bills regardless of maturity and bonds with fewer than 3 months to maturity, but includes almost all straight notes and bonds. In addition, a high number of assets allows much more complex yield curve shapes.

The main factor underlying the different security sample composition is liquidity, which can have an important impact on prices. "On-the-run" issues often trade at a premium with respect to the remainder of the outstanding issues.[7] A just-issued bond concentrates most of the trading volume because most investors and fund managers are trying to allocate or distribute this new asset in their portfolios or among their clients. Specialness in the repo market may also cause on-the-run securities to trade at a premium.[8] Treasury supply and demand and other external factors such as market-level illiquidity, monetary policy decisions and macroeconomic shocks can affect the spread between on-the-run and off-the-run government bonds.[9] In this sense, the DoT and the FRB take opposing positions regarding liquidity, i.e., the DoT considers only "on-the-run" issues, whereas the FRB includes "second-off-the-run" bonds or older. F082 includes all of the securities. Additionally, the Treasury bills yields can exhibit idiosyncratic variations that seem distinct from the other longer-dated segments of the Treasury markets.

3 Isolating the effect of the data choice

In this section we study the effect of the set of securities used by each zero-coupon interest data set on several crucial variables for a variety of financial purposes. As mentioned above, each data set (DoT, FRB or F082) employed different assets and

[7] See, e.g., Sarig and Warga (1989), Warga (1992), Fleming (2000), Amihud and Medelson (1991), Krishnamurthy (2002), Sack and Elsasser (2004), Goldreich, Hanke and Nath (2005), Díaz et al. (2006), Vayanos and Weill (2008), Pasquariello and Vega (2009), Díaz et al. (2011), Graveline and McBrady (2011) and Fontaine and García (2012), and Díaz and Escribano (2017).

[8] Intermediaries in the repo market require Treasuries as collateral; for them, on-the-run bonds are appealing securities. These securities may trade on special, i.e., it can be used as collateral to borrow money at a rate below the prevailing general repo rate either because it is more liquid or because of its scarcity caused by a limited supply or a short squeeze.

[9] See, e.g., Duffie (1996), Jordan and Jordan (1997), Krishnamurthy (2002), Fleming (2003), Longstaff (2004), Cherian et al. (2004), Krishnamurthy and Vissing-Jorgensen (2012) and Barnejee and Graveline (2013).

prices to estimate the yield curve. The impact of this factor on the yield curve cannot be isolated from other sources of discrepancy such as the alternative methodologies employed by each data provider. To solve this problem, we propose an indirect comparison method to isolate the effect of the input choice made by these providers on the resulting spot interest-rate estimates. First, we generate our own yield curve estimates using a common and popular fitting method, the Svensson's (1994) approach, but using different baskets of assets that attempt to mimic those used by each yield curve provider. This approach also avoids the fact that each data set analyzed use different bond prices: "end-of-the-day" prices (FRB), "close of business" bid yields-to-maturity (DoT), or what Bloomberg calls "generic prices" (F082). Second, we check if there are also significant differences on the pricing of traded Treasury securities, the resulting volatilities and the pricing of callable bonds, the correlations among forward rates, and a simple expectations hypothesis test.

3.1 Mimicking sample compositions

We obtain intraday US Treasury security quotes and trades for all issues during the period between January 1996 and December 2006 (2864 trading days) from the GovPX database. The quality and quantity of information provided by GovPX progressively fades over time, so we decide to truncate our data sample at the end of 2006. Further technical details of our preparation of these data appear in the "Appendix".

From the daily prices we extract from the GovPX database, we propose four alternative combinations of securities that replicate those used by the external providers of yield curves. We analyze the impact of asset selection on the resulting yield curve by mimicking the inputs used by our three datasets and obtaining our own estimates of zero-coupon rates using a single method.

Table 2 reports detailed information about the composition of our dataset from GovPX asset prices. "Sample A"—the "full sample"—refers to the sample that includes all bills, notes and bonds that meet the requirement described above. This sample can be considered a benchmark or basic sample. Excluding the on-the-run and the first off-the run assets of each maturity from Sample A gives us "Sample B," which can be considered an intermediate case. Separately, "Sample C" approximately mimics the FRB sample composition by excluding all the bills and the on-the-run and first off-the-run securities. Finally, "Sample D" mimics the DoT sample composition, which only includes the on-the-run bills, notes and bonds for eleven maturities: four maturities of the most recently auctioned bills (4-, 13-, 26- and 52 weeks), six maturities of just-issued notes and bonds (2-, 3-, 5-, 7-, 10-, and 30 years) and the composite rate in the 20-year maturity range.

Given the above commented on/off-the-run phenomenon, one would certainly expect the yield curve implied by Sample D (on-the-run bonds) to be different from those implied by Samples A and B (which are mostly populated by off-the-run bonds). Also, the maturity spectrum, the number of securities included in the fitting process, and the relative number of bonds with longest maturities can imply a different behavior of Sample D. In addition, the term structures implied by Samples B and C (same as

Table 2 Composition of our dataset from GovPx asset prices

Year	# observations per day				Days per year	Maturity longest bond
	(A)	(B)	(C)	(D)		
1996	151.6	137.6	115.2	11.0	259	29.8
1997	150.1	136.2	114.0	11.0	261	29.8
1998	146.7	132.7	109.4	11.0	261	29.7
1999	134.4	120.4	96.8	11.0	261	29.8
2000	120.0	106.1	84.6	11.0	260	29.8
2001	114.6	100.6	79.6	11.0	258	29.5
2002	113.3	99.3	76.7	11.0	261	28.6
2003	115.2	101.2	78.6	11.0	261	27.6
2004	123.8	109.8	87.2	11.0	262	26.6
2005	129.6	115.7	93.7	11.0	260	25.6
2006	157.0	143.4	122.0	11.0	260	28.6
Avg	132.4	118.4	96.1	11.0	260	28.7

This table shows the average number of observations per day, the number of trading days in the year and the average maturity of the longest bond included in the daily estimation. The dataset includes intraday US Treasury security quotes and trades for all issues during the period between January 1996 and December 2006 (2864 trading days)

(A) Full sample: bills, notes, and bonds

(B) Full sample excluding on-the-run and first off-the-run notes and bonds

(C) Mimicking FRB (excluding bills, and on-the-run, first off-the-run notes and bonds)

(D) Mimicking DoT (only on-the-run securities) (11 maturities)[a]

All of the samples exclude "when-issued" and cash management transactions, trades and quotes related to callable bonds, and TIPS, along with outliers (usual filters are applied)

[a]Four maturities of the most recently auctioned bills (4-, 13-, 26-, and 52-weeks), six maturities of just-issued bonds and notes (2-, 3-, 5-, 7-, 10-, and 30-years), and the composite rate in the 20-year maturity range

Sample B except all bills are excluded) might be different because the idiosyncrasy of the Treasury bill segment.

3.2 Term structure specification

As mentioned above, we cannot replicate the fitting methods of two of the yield curve databases because we do not have enough information about the DoT and Bloomberg approaches. Thus, we apply Svensson's (1994) method to the four different sets of assets.[10] It should be noted that this method is the one used in FRB data set; however, there are relevant differences between the methodology that Gürkaynak et al. (2007) employ for this fitting approach and the methodology applied in this paper.

[10] According to BIS (2005), nine out of thirteen central banks currently use either the Nelson and Siegel (1987) model or its extended version suggested by Svensson (1994) to estimate the term structure of interest rates. One of the exceptions is the United States, which applies a "smoothing splines" method.

The estimation of the spot rates on a given date involves finding a functional form that approximates the theoretical discount function, $D(t)$, and replicating a set of bond prices at a given instant as accurately as possible:

$$P_k = \sum_{T=T_1^k}^{T_n^k} C_T^k \cdot D(T, \bar{b}) + \varepsilon_k \quad k = 1, 2, \ldots, m \tag{1}$$

where P_k is the price of bond k, C_T^k denote the cash flows (coupon and principal payments) generated by bond k, $D(T, \overline{b})$ is the discount function that we want to approximate and that depends on a vector of parameters $\overline{b}$, ε_k is an error term, and m is the number of different assets included in the estimation.

The Svensson (1994) "two hump" model provides a reliable fit. This parsimonious approach imposes a functional form for the instantaneous forward rate, which are governed by six parameters:

$$f_T = \beta_0 + \beta_1 \exp\left(-\frac{T}{\tau_1}\right) + \beta_2 \frac{T}{\tau_1} \exp\left(-\frac{T}{\tau_1}\right) + \beta_3 \frac{T}{\tau_2} \exp\left(-\frac{T}{\tau_2}\right) \tag{2}$$

where T is the term to maturity and (β_0, β_1, β_2, τ_1, β_3, τ_2) is the set of parameters to be estimated.

The last two parameters make the difference with respect to the Nelson and Siegel's (1987) method. The latter term allows a second "hump" in the forward rate curve and provides a better fit to the convex shape of the yield curve at the long end.

Numerous authors have interpreted the first three parameters in the Nelson and Siegel model—β_0, β_1, and β_2—as the specific factors that drive the yield curve: level, slope and curvature. β_0 is related to the long-run level of interest rates (the limit of the spot rate function as maturity approaches infinity); β_1 is regarded as the long-to-short-term spread (limit of the spot rate as maturity approaches zero minus β_0) and β_2 is related to the yield curve curvature because its changes almost exclusively affect medium-term interest rates. The parameter τ_1 determines the maturity at which the curvature factor β_2 is maximized and the speed of the exponential decay of the slope factor β_1.

With respect to the residuals ε_k, we discard assuming homoskedasticity and applying the ordinary least squares method to estimate the parameters. This assumption would imply assigning the same importance to errors in the price of all bonds, which means that we very heavily penalize errors in the yields of long-term bonds. Therefore, assuming homoskedasticity implies forcing the adjustment at the long end of the yield curve at the cost of relaxing the adjustment of the curve for short maturities.

To address this problem, some authors have suggested penalizing the valuation errors of short-term bonds; in particular, recommendations usually propose correcting the variance of the error term, making it proportional to the bond duration:

$$VAR\left[\varepsilon_k^2\right] = \left(\frac{\partial P_k}{\partial y_k}\right)^2 \cdot \sigma^2 = \left(\frac{D_k \cdot P_k}{1 + y_k}\right)^2 \cdot \sigma^2 \tag{3}$$

where D_k is the k-bond duration, y_k is its yield to maturity, and P_k is its price. Next, the model is adjusted using generalized least squares (GLS). This is the approach we follow in this article. In this way, we force the adjustment of the short-term interest rates. However, this adjustment is not free: it implies that we relax the adjustment of long-term interest rates.

Finally, we estimate the six parameters of the Svensson model simultaneously using nonlinear optimization techniques.[11] Although the time series of both τ_1 and τ_2 are volatile, the numerical stability of the other four parameters is remarkable.

4 The impact of the data choice

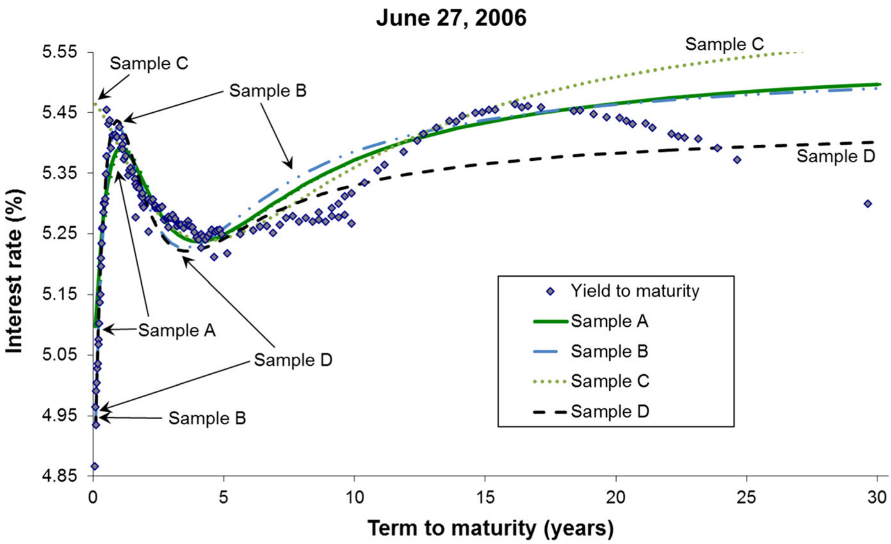
Before proceeding to our statistical analysis, we present two examples to illustrate the problems derived from using different baskets of assets to estimate the yield curve. Thus, on June 27, 2006, the shape of the observed yields to maturity included in the sample is complex enough to show differences in fitting flexibility. The second example, which corresponds to October 10, 1999, clearly highlights the impact of liquidity.

Panels A and B of Fig. 1 depict four yield curve estimations using the same fitting technique but different baskets of assets. This allows us to illustrate the impact of the asset sample composition. In the first example (June 27, 2006, Panel A), both the full sample (Sample A) and the sample excluding on-the-run and first off-the-run bonds (Sample B) produce parallel term structure estimates with small differences. The only remarkable difference is that excluding on-the-run seems to closer fit the data for around 1 year of maturity. The 1-year on-the-run bill included in Sample A pulls down the spot-rate estimations at the short end of the yield curve. Excluding bills and on-the-run bonds (Sample C) also accurately fits 1-year maturity rates but performs poorly for the shortest maturities. Finally, including only on-the-run assets (Sample D) produces yield curves that are clearly conditioned by the sample composition. The maturity of six of the eleven considered on-the-run securities are shorter than 3 years. This yield curve shows a good fitting for short maturities. Meanwhile, the resulting zero-coupon rates for maturities longer than 10 years differ considerably from the actual data.

The second example (Panel B) illustrates the impact of liquidity involved in the assets considered by the different samples. Excluding on-the-run securities (Samples B and C) seems to slightly pull up the yield curve compared with Sample A for maturities longer than 4 years. The absence of bills in Sample C (mimicking FRB) again implies a poor fitting at the short end of the yield curve. Conversely, Sample D (mimicking DoT), which only includes on-the-run bills and bonds, provides a good fitting at the short end of the curve but poor performance for most of the remaining maturities.

[11] Even to apply the Nelson and Siegel (1987) technique, Diebold and Li (2006) propose to fix τ_1 to simplify the estimation of the remaining three parameters with the advantage of providing better numerical stability.

(A) June 27, 2006



(B) October 8, 1999

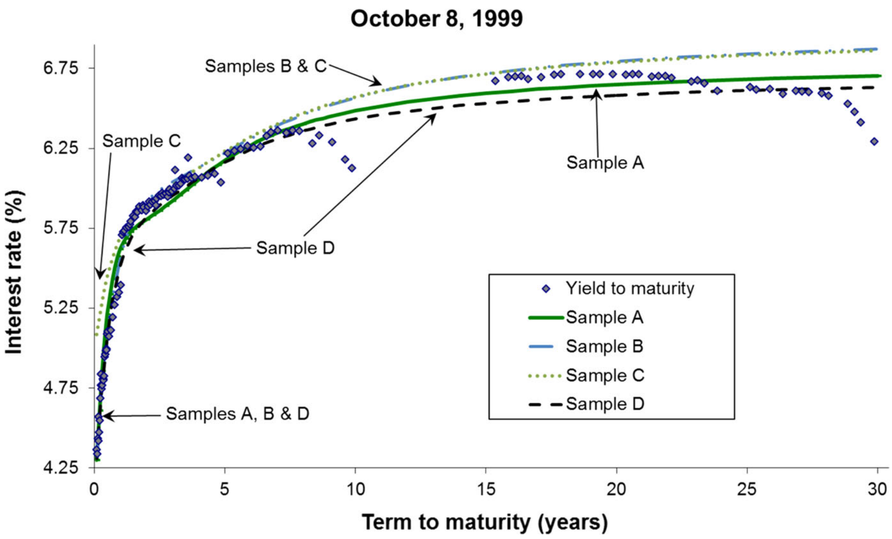


Fig. 1 Impact of the basket of assets on the spot rates. Yield curve estimates applying the weighted Svensson to the four different datasets from GovPx asset prices (Sample A: full sample (bills, notes, and bonds); Sample B: full sample excluding on-the-run and first off-the-run; Sample C: mimicking FRB, excluding bills, on-the-run, first off-the-run; Sample D: mimicking DoT, only on-the-run, 11 maturities. *Note* The points represent the yields to maturity of all of the non-callable traded Treasury securities (Sample A or the full sample)

Table 3 Estimated parameters in yield curve estimates using Svensson (1994)

Basket of assets	β_0	β_1	β_2	τ_1	β_3	τ_2
Mean						
Sample A (full sample)	6.18	− 0.52	0.18	1.33	0.25	1.63
Sample B (excluding OTR and 1st OffTR)	6.21	− 0.48	0.16	1.25	0.23	1.53
Sample C (FRB: excluding bill, OTR and 1st OffTR)	6.24	− 0.86	0.37	1.78	0.42	2.04
Sample D (DoT: excluding OffTR)	6.04	− 0.49	0.17	1.20	0.23	1.44
Standard deviation						
Sample A (full sample)	0.69	0.26	0.16	0.62	0.12	0.91
Sample B (excluding OTR and 1st OffTR)	0.68	0.25	0.16	0.42	0.12	0.38
Sample C (FRB: excluding bill, OTR and 1st OffTR)	0.68	0.24	0.13	0.89	0.12	1.86
Sample D (DoT: excluding OffTR)	0.67	0.25	0.16	0.57	0.12	0.54

This table reports averages of estimated primary parameters of the weighted Svensson (1994) model from GovPx asset prices. We consider 2864 daily estimates of the term structure of interest rates for each basket of assets during the period between January 1996 and December 2006. The four samples are Sample A (full sample, including all bills and straight notes and bonds); Sample B (full sample excluding on-the-run and first off-the-run); Sample C (mimicking FRB, excluding bills, on-the-run, first off-the-run); and Sample D (mimicking DoT, only on-the-run, 11 maturities)

4.1 The impact on the spot rates

4.1.1 Yield curve parameters

Table 3 presents descriptive statistics of the estimated parameters in the Svensson model fitting to all samples. On average, Samples A, B and D provide similar results for the six estimated parameters. Recall that the parameters can be understood as the very long interest rate (β_0), the slope (β_1) and the curvature (β_2 and β_3) of the yield curve, and a measure of the rate at which the short- and medium-term components decay to zero (τ_1 and τ_2). The result for Sample C is particularly interesting. All of the parameters, with the exception of β_0, show remarkably different average values in Sample C. Considering that excluding bills is the only difference between Samples B and C, the lack of these short-term securities has relevant implications. The shape of the entire yield curve—i.e., the slope, the curvature and the position of possible humps—depends on the assets considered at the very beginning of the maturity spectrum. In addition, Sample C tends to overestimate the interest rates of the shortest maturities because this sample ignores bills that are more liquid than old bonds with similar remaining maturity, which therefore trade at lower yields to maturity.

4.1.2 Pricing traded treasury securities

A direct way to check the accuracy of the estimated yield curves is to use them to price Treasury bills, notes and bonds. Comparing the actual traded price with the estimated price discounting cash flows with the appropriate zero-coupon rate allows pricing errors to be obtained. Because we use the same fitting technique to estimate the four yield curve samples, the basket of assets is the only difference among them.

Table 4 shows the root mean squared error (RMSE) and the mean absolute error (MAE), which is less easily influenced by extreme observations, obtained from pricing eighteen securities the first working day of each month in the sample period. For each date, we select the currently on-the-run asset, i.e., the just issued, the first off-the-run and the second off-the-run. Specially, we consider the 6- and 12-month on-the-run bills, and the 2-, 5-, 10-, and 30-year on-the-run notes and bonds. Obviously, the term to maturity of the first off-the-run is shorter than that of the on-the-run asset, and likewise the term of the second off-the-run is shorter than that of the first off-the-run. The analyzed securities change for each date.

Sample C, which considers neither bills nor on-the-run bonds, provides the worst results in the short- and long-term tranches, i.e., when pricing bills and 10- and 30-year bonds. On the other hand, this sample performs best when pricing 5-year bonds and 2-year off-the-run bonds. The inclusion of the bills in the asset basket, i.e., changing from sample C to sample B, considerably improves the adjustment in the short term (12 months or less) but worsens the adjustment for all other maturities. The inclusion in the basket of on-the-run and first-off-the-run assets (Sample A) slightly improves the fitting for all maturities relative to the results of sample B, especially for long maturities. Finally, the sample that only includes 11 on-the-run assets obtains modest results in the short term, poor results for 2-year off-the-run and all 5-year bonds but is by far the one that provides the best performance for long terms. The differences in accuracy when valuing 30-year bonds are remarkable compared to the other samples. Samples A, B and C underestimate the price of the 30-year bonds for all the percentiles, especially in the case of the on-the-run bond. As shown in Table 3, the smallest value of β_0 in Sample D provides lower long-term zero-coupon interest rates, which allows bond prices to be obtained closer to the real ones. This result is unsurprising given that Sample D, whose composition mimics that of the DoT, gives a relevant weight to the long term, with three of the considered eleven securities with terms to maturity of 10 years or more (i.e., 10-, 20- and 30-year bonds). This is an interesting result on the valuation of liabilities of pension funds.

We can conclude highlighting several results. First, the performance in terms of pricing errors of the different yield curves depends on the maturity range considered. Without there being a sample that beats all the others for the whole spectrum of terms, we can highlight that sample D, with only 11 observations and a very reduced computational effort at the time of its calculation, provides reasonable results in the short term and very good results for the on-the-run 2-year note and maturities of 10 and 30 years. Second, performance hardly varies when pricing assets with different status, except in the case of 2-year note. The special impact of liquidity on the 2-year on-the-run note complicates the correct pricing. For instance, sample D is the one that best values this on-the-run note but the one that worst values the off-the-run notes of that maturity.

Table 4 Summary statistics of pricing errors for eighteen assets from the yield curve estimates using Svensson (1994)

Basket of assets	Avg. RMSE when pricing on-the-run assets							Avg. MAE when pricing on-the-run assets						
	RMSE	6-month	12-month	2-year	5-year	10-year	30-year	MAE	6-month	12-month	2-year	5-year	10-year	30-year
Panel A: Pricing on-the-run bills (6- and 12-month), notes and bonds (2-, 5-, 10-, and 30-year)														
Sample A	0.697	0.001	0.007	0.013	0.085	2.619	11.393	0.771	0.019	0.062	0.083	0.211	1.424	3.124
Sample B	0.785	0.001	0.008	0.013	0.091	3.032	12.654	0.821	0.020	0.067	0.086	0.220	1.515	3.300
Sample C	0.795	0.006	0.017	0.010	0.079	2.933	12.832	0.829	0.056	0.091	0.073	0.209	1.474	3.346
Sample D	0.218	0.002	0.008	0.008	0.273	1.147	3.469	0.416	0.021	0.058	0.059	0.249	0.920	1.675
Basket of assets	Avg. RMSE when pricing first off-the-run assets							Avg. MAE when pricing first off-the-run assets						
	RMSE	6-month	12-month	2-year	5-year	10-year	30-year	MAE	6-month	12-month	2-year	5-year	10-year	30-year
Panel B: Pricing first off-the-run bills (6- and 12-month), notes and bonds (2-, 5-, 10-, and 30-year)														
Sample A	0.232	0.000	0.005	0.008	0.041	1.109	3.717	0.428	0.017	0.051	0.060	0.133	0.908	1.751
Sample B	0.280	0.000	0.006	0.008	0.037	1.335	4.419	0.476	0.017	0.055	0.060	0.134	0.996	1.930
Sample C	0.287	0.006	0.013	0.005	0.033	1.239	4.565	0.485	0.056	0.080	0.048	0.120	0.950	1.975
Sample D	0.042	0.002	0.008	0.011	0.349	0.447	0.502	0.141	0.022	0.056	0.077	0.286	0.494	0.504
Basket of assets	Avg. RMSE when pricing first off-the-run assets							Avg. MAE when pricing first off-the-run assets						
	RMSE	6-month	12-month	2-year	5-year	10-year	30-year	MAE	6-month	12-month	2-year	5-year	10-year	30-year
Panel C: Pricing second off-the-run bills (6- and 12-month), notes and bonds (2-, 5-, 10-, and 30-year)														
Sample A	0.121	0.000	0.003	0.006	0.043	0.673	1.909	0.313	0.015	0.038	0.054	0.138	0.709	1.266
Sample B	0.154	0.000	0.003	0.006	0.035	0.830	2.419	0.358	0.015	0.041	0.054	0.130	0.791	1.442
Sample C	0.160	0.006	0.010	0.004	0.034	0.743	2.545	0.367	0.055	0.069	0.045	0.124	0.742	1.485
Sample D	0.030	0.002	0.006	0.010	0.380	0.249	0.380	0.096	0.021	0.046	0.074	0.316	0.366	0.460

These tables report the root mean squared errors (RMSE) and mean absolute errors (MAE) in pricing the on-the-run, first off-the-run, and second off-the-run asset for several maturities (6- and 12-month, 2-, 5-, 10-, and 30-year) from the yield curve estimates on the first working day of each month during the sample period. Only in-sample pricing errors are computed. The four samples are Sample A (full sample, including all bills and straight notes and bonds); Sample B (full sample excluding on-the-run and first off-the-run); Sample C (mimicking FRB, excluding bills, on-the-run, first off-the-run); and Sample D (mimicking DoT, only on-the-run, 11 maturities)

4.2 The impact on the volatility term structure

4.2.1 Comparing volatility term structures

Previous analyses in the literature focus on the ability of alternative zero-coupon yield estimates using different fitting techniques to replicate bond market prices or to adjust arbitrage-free dynamic term structure models. Instead, we study potential implications of the set of bond prices used in the yield curve construction on the estimates of other variables that imply using the time series of interest rates: the volatility term structure (VTS) and some correlations between pairs of forward rates. Both variables play a key role in valuation models of interest rate derivatives and risk management calculations.

To estimate interest rate volatilities, we use two alternative methods. First, we calculate the simple standard deviations using a 3-month rolling window with a daily frequency from log-differences of the value of the spot rate for eleven maturities ranging from 1 month to 30 years. We refer to the resulting annualized volatilities as "historical volatilities." Second, we consider different specifications of the well-known family of conditional volatility models. Based on the Schwartz Information Criterion and the Akaike Information Criterion (SIC and AIC, respectively), we choose the EGARCH(1,1) model proposed by Nelson (1991) which allows for asymmetric impacts of innovations.[12]

Table 5 summarizes some statistical results of the VTS estimations. Both the historical method (panel A) and the conditional method (panel B) produce quite similar averages. Both panels also show decreasing standard deviation values of the realized volatility as maturity increases. At the short end of the yield curve, the distribution of realized volatility is fat-tailed and slightly skewed. These outcomes are robust and independent of the method used to compute volatility. However, the VTS obtained as historical volatilities provide better statistical properties than those computed from the conditional model. Main differences between both methods are observed for the higher-order moments of the volatility distribution in the two shortest maturities (1- and 3-month). The conditional method provides extreme large positive deviations in the short end of the VTS.

In Table 6, we summarize the statistical significance of the difference observed in the panel A of Table 5. The volatilities are computed by annualized standard deviation using 3-month non-overlapping windows. A "X/Y ratio" at time t is calculated dividing the volatility level from the dataset X at time t by the corresponding volatility from the Y dataset at time t. The table shows average X/Y ratio, the percentage of times the ratio is higher than the unity, a standard *t* test, a common non-parametric sign test (binomial test) and a signed-ranked test where the null hypothesis is that the X/Y ratio is equal to unity, i.e., there is no significant difference between the volatility level of both the X and the Y datasets. We also compute the test statistic proposed by Levene (1960) to test the null hypothesis of equal variances in each pair of volatility series.

There are at least four noteworthy results shown in Tables 5 and 6. First, dropping bills and on-the-run notes and bonds (Sample C) from the entire sample (Sample A)

[12] Andersen and Benzoni (2007) have documented that EGARCH representation for conditional yield volatility provides a convenient and successful parsimonious model for the conditional heteroskedasticity in these series.

Table 5 Descriptive statistics of the volatility term structure of yield curve estimates using Svensson (1994)

t	Mean				Std.dev.				Median				Skewness				Excess kurtosis			
	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Panel A: Descriptive statistics of historical volatility from standard deviations																				
1 m	0.57	0.57	0.35	0.81	0.14	0.13	0.29	0.16	0.58	0.57	0.26	0.79	0.24	0.54	1.50	2.72	− 1.14	− 0.55	2.29	11.24
3-month	0.37	0.37	0.27	0.44	0.10	0.10	0.20	0.11	0.32	0.33	0.20	0.41	1.06	1.09	1.21	1.53	0.27	0.33	0.92	2.91
6-month	0.24	0.24	0.23	0.28	0.13	0.13	0.16	0.18	0.18	0.18	0.15	0.21	0.93	0.89	0.95	1.68	− 0.28	− 0.40	− 0.26	3.83
1-year	0.25	0.25	0.25	0.28	0.18	0.18	0.19	0.20	0.16	0.15	0.15	0.19	0.99	0.98	0.96	1.60	− 0.54	− 0.58	− 0.64	2.76
2-year	0.28	0.28	0.27	0.29	0.19	0.19	0.19	0.17	0.19	0.18	0.18	0.19	1.09	1.08	1.13	0.93	− 0.21	− 0.21	− 0.07	− 0.57
3-year	0.27	0.26	0.26	0.28	0.16	0.16	0.16	0.14	0.20	0.20	0.19	0.21	1.13	1.14	1.17	1.02	0.07	0.13	0.19	− 0.10
5-year	0.23	0.23	0.23	0.25	0.11	0.10	0.10	0.10	0.20	0.19	0.19	0.21	1.10	1.09	1.09	0.95	0.30	0.32	0.28	− 0.05
7-year	0.21	0.20	0.20	0.22	0.08	0.07	0.07	0.08	0.19	0.18	0.18	0.20	0.99	0.98	0.98	0.84	0.28	0.28	0.26	− 0.07
10y	0.18	0.17	0.17	0.18	0.05	0.05	0.05	0.05	0.17	0.16	0.16	0.17	0.86	0.87	0.87	0.67	0.24	0.26	0.28	− 0.31
20y	0.15	0.15	0.14	0.16	0.03	0.03	0.03	0.03	0.14	0.14	0.14	0.16	0.71	0.69	0.75	1.53	0.37	0.39	0.50	4.33
30y	0.16	0.14	0.14	0.18	0.05	0.03	0.03	0.06	0.14	0.14	0.14	0.17	1.86	0.58	0.65	3.04	3.26	0.07	0.24	12.23
Panel B: Conditional volatility from EGARCH(1,1)																				
1-month	0.53	0.53	0.35	0.77	0.34	0.36	0.31	0.49	0.44	0.43	0.24	0.62	16.49	19.60	2.83	18.26	444.08	666.40	15.19	528.94
3-month	0.36	0.35	0.27	0.43	0.34	0.21	0.21	0.29	0.29	0.30	0.19	0.36	31.47	21.60	2.17	26.96	1249.35	734.17	7.86	1038.21
6-month	0.24	0.24	0.24	0.28	0.13	0.13	0.17	0.16	0.18	0.18	0.16	0.22	1.15	1.11	1.38	1.25	0.82	0.70	1.77	1.82
1-year	0.26	0.25	0.26	0.29	0.18	0.18	0.19	0.18	0.16	0.15	0.16	0.20	1.05	1.04	1.02	1.42	− 0.30	− 0.33	− 0.39	2.30
2-year	0.29	0.28	0.28	0.29	0.19	0.18	0.19	0.17	0.19	0.19	0.18	0.20	1.14	1.13	1.17	1.06	0.03	− 0.01	0.10	− 0.03
3-year	0.27	0.27	0.26	0.28	0.15	0.15	0.15	0.14	0.20	0.20	0.19	0.22	1.16	1.18	1.20	1.10	0.21	0.27	0.32	0.20
5-year	0.24	0.23	0.23	0.25	0.10	0.10	0.10	0.09	0.20	0.19	0.19	0.22	1.10	1.11	1.11	0.97	0.19	0.31	0.31	0.00
7-year	0.21	0.20	0.20	0.22	0.07	0.07	0.07	0.07	0.19	0.18	0.18	0.20	0.98	0.99	0.96	0.83	0.07	0.17	0.12	− 0.23

Table 5 continued

t	Mean				Std.dev.				Median				Skewness				Excess kurtosis			
	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
10-year	0.18	0.17	0.17	0.18	0.05	0.04	0.04	0.05	0.17	0.17	0.17	0.17	0.80	0.81	0.78	0.66	− 0.10	− 0.07	− 0.09	− 0.20
20-year	0.15	0.15	0.14	0.16	0.02	0.02	0.02	0.02	0.15	0.14	0.14	0.16	0.48	0.39	0.53	1.13	− 0.33	− 0.39	− 0.21	2.85
30-year	0.16	0.14	0.14	0.18	0.03	0.02	0.02	0.02	0.15	0.14	0.14	0.18	1.35	0.40	0.66	7.35	1.85	− 0.31	0.20	146.27

Panel A reports daily averages of the volatility term structure (VTS) computed by two alternative methods: annualized simple standard deviation measures using 30-day rolling windows from the log-difference of the value of the spot rates (“historical volatilities”), and an EGARCH(1,1) model. As input, we use our zero-coupon interest rate estimates using four different set of bond prices from GovPx (Sample A: full sample (bills, notes, and bonds); Sample B: full sample excluding on-the-run and first off-the-run; Sample C: mimicking FRB, excluding bills, on-the-run, first off-the-run; Sample D: mimicking DoT, only on-the-run, 11 maturities. The data is daily and cover the period from Jan. 1996 to Dec. 2006

Table 6 Test results of the volatility term structure of yield curve estimates using Svensson (1994)

t	Ratio A/C						Ratio B/C					
	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test
1-month	1.93	81	5.98***	4.19***	5.22***	14.27***	1.72	83	5.59***	4.47***	5.38***	15.96***
3-month	1.65	77	5.98***	3.61***	4.76***	16.11***	1.45	79	6.15***	3.90***	5.08***	16.50***
6-month	1.14	63	2.84***	1.58	2.47**	4.09**	1.13	73	3.91***	3.03***	3.50***	3.29*
1-year	1.00	50	0.89	− 0.14	0.05	0.40	1.00	50	0.10	− 0.14	0.15	0.22
2-year	1.02	56	1.17	0.72	0.23	0.13	1.00	56	2.08**	0.72	1.73*	0.03
3-year	1.01	50	1.18	− 0.14	0.08	0.06	1.01	67	3.65***	3.14***	3.28***	0.00
5-year	1.01	54	1.10	0.43	0.10	0.03	1.01	67	1.50	2.17**	2.33**	0.03
7-year	1.01	52	1.09	0.14	− 0.01	0.07	1.01	69	0.90	2.45**	1.95*	0.02
10-year	1.00	52	0.94	0.14	0.09	0.14	1.00	50	− 0.34	− 0.14	0.15	0.02
20-year	1.00	52	1.63	0.14	1.04	0.03	1.00	69	3.19***	2.45**	2.90***	0.00
30-year	1.04	65	2.72***	1.88*	2.29**	3.41*	1.02	71	0.99	2.74***	2.57**	0.09

t	Ratio C/D					
	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test
1-month	0.75	6	− 9.24***	5.92***	5.82***	15.15***
3-month	0.85	6	− 8.90***	5.92***	5.85***	14.83***
6-month	0.97	38	− 3.63***	1.59	3.00***	0.03
1-year	0.92	38	− 4.18***	1.59	3.13***	0.02
2-year	0.95	38	− 3.79***	1.59	2.99***	0.58
3-year	0.97	35	− 4.15***	1.88*	3.45***	0.25

Table 6 continued

t	Ratio C/D					
	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test
5-year	0.95	25	– 5.56***	3.32***	4.83***	0.01
7-year	0.95	19	– 6.13***	4.19***	5.09***	0.13
10-year	0.96	19	– 5.74***	4.19***	5.21***	0.15
20-year	0.95	29	– 5.42***	2.74***	4.11***	0.47
30-year	0.88	29	– 5.95***	2.74***	4.42***	3.61*

This table reports the test results for different pairs of volatility term structure (VTS) samples. Each VTS is computed as annualized simple standard deviation measures using 30-day rolling windows from the log-difference of the value of the spot rates ("historical volatilities"). As the input, we use zero-coupon interest rates from our own yield curve estimates, applying the weighted Svensson to the four different datasets from GovPx asset prices (Sample A: full sample (bills, notes, and bonds); Sample B: full sample excluding on-the-run and first off-the-run; Sample C: mimicking FRB, excluding bills, on-the-run, first off-the-run; Sample D: mimicking DoT, only on-the-run, 11 maturities. We consider 2864 daily estimates of the term structure of interest rates for each basket of assets during the period between January 1996 and December 2006

The X/Y ratio is defined as the mean volatility for the X dataset divided by the corresponding mean for the Y dataset. The hypothesis that the mean of the X/Y ratio is equal to unity is tested using a standard *t*-test. A standard sign test is used to test that the proportion of X/Y ratios higher than unity is equal to the proportion of ratios lower than unity, which is equal to 0.5. The Levene statistic is used to test the equality of variances between X and Y. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively, in a two-tailed test causes the resulting volatility of all interest rates to decrease. Considering these assets provides a closer fit of the short end of the yield curve (see Table 4), but increases the mean of the 1-month volatility by as much as 93%. This is caused by the fact that incorporating bills and on-the run bonds increases the variability of assets with similar terms to maturity but different yields because of differences in liquidity. This is an extra "noise" that can increase the variance of the estimated zero-coupon rates. The result is statistically significant at the short and long ends of the term structure.

Second, average values for Sample B (including bills) and Sample C (excluding bills) reported in Table 5 are quite similar for all the maturities higher than 6 months. However, including bills provides statistically significant higher values of the volatility than excluding bills for all the maturities (see central panel in Table 6). There are two curious exceptions, the null hypothesis of equal means and equal medians between including/excluding bills cannot be rejected at the 1- and 10-year maturities.

Another interesting difference appears in Sample D, which produces much higher estimates of volatility than any other sample at both ends of the yield curve. The different behavior of the volatility level of only including on-the-run assets is significant from a statistical point of view for the entire maturity spectrum. Some of these differences are truly overwhelming. The reason for these results is the small number of long-term assets included in that Sample D. This generates very unstable estimates of the zero-coupon rates and so the time series of yields present higher volatilities. At the same time the higher volatility of the time series of short-term interest rates can be caused by the importance of short-term instruments in this sample. This provokes a good fit in this part of the yield curve, capturing the actual behavior of short rates. Finally, the variance of volatility for the shortest maturities and for the 30-year maturity is significantly different among the four interest rate datasets.

Overall, we can see that the basket of assets eventually used to estimate zero-coupon rates implies not only a better or worse fitting in some tranches of the yield curve but also important consequences in their volatility estimates. The dissimilar behavior observed in four datasets may have relevant implications. In this sense, Corsi et al. (2008) highlights the importance of the time-varying features of the realized volatility for several applications, such as Value-at-Risk calculations, or pricing or assessing the risk of volatility derivatives.

4.2.2 Pricing callable default-free bonds

In this section, we analyze the implications for the pricing of interest rate derivatives of the use of different yield curves depending on the basket of data used in their estimation. We calculate the price and delta of call options written in theoretical callable bonds from the four yield curves. We consider default-free bonds to avoid the iterations between interest risk and default risk. To value the callable bonds, we adjust the well-known model of Black et al. (1990) for the first working day of March and November during the sample period. This model is one of the most used by the industry. It adjusts the parameters to exactly replicate the term structures of interest rates and volatilities. In the calibration process for each date, we use zero coupon rates for monthly terms up to 30 years and the associated volatilities for each of these terms. We consider the annualized volatilities referred to "historical volatilities" in the previous section.[13]

From the information provided by the Fixed Income Securities Database (FISD) during the sample period, we study some of the most frequent call structures among the US corporate bonds. We propose eight callable bonds with terms to maturity between 1 year and 3 months and 30 years, and with one or two embedded call options (1.25-year bond with a call date 3 months hence, 3-year bond callable after 1.25 years, 10-year bond callable at 2-year, 20-year bond callable at 5- and 6-year, 20-year bond callable at 18- and 19-year, 30-year bond with two call dates 3 and 4 years hence, 30-year bond callable at 5-year, and 30-year bond callable at 10-year). As for the coupon rate, we set it at 5% payable semi-annually, which is close to the average coupon rate during the sample period, which is 5.35%. The bond has a face value of $100 and can be redeemed by the issuer at par value of 100 at the callable date or dates.

Using the four yield curve datasets, we obtain the price and yield-to-maturity (YTM) of the noncallable bond by discounting the semiannual coupons and principal. The price of the callable bond is computed from the calibrated BDT model. By subtracting the callable bond price from the price of the corresponding noncallable bond, we infer the value of the call option implicit in the price of the callable bond. We also compute the impact of the call on the YTM. Finally, the call delta is the ratio between the change in price of the call option and the change in price of the underlying bond.

For each date in which the BDT model is calibrated for each database, we calculate the five variables: price and YTM of the noncallable bond, call impact on the price and YTM of the callable bond, and call delta. For each one of these variables V, we calculate the relative variation $[(V_{jt} - AvgV_t)/AvgV_t]$ of the value obtained for each sample j on date t (V_{jt}) with respect to the average value of the variable for the four samples on date t $(AvgV_t = \Sigma_j\ V_{jt})$. Table 7 depicts the average of these relative variations along the sample period. These figures reflect the pricing differences for each sample with respect to the average of the four databases.

There are no significant differences when pricing noncallable bonds. In the case of 30-year bonds, Sample D provides prices 1.48% higher than the average of the four databases, which corroborates the results obtained in Sect. 4.1.2. However, in the valuation of the call option, Samples A and D generally show values above the mean, while Samples B and C show them below. The two samples including on-the-run assets in their baskets (Samples A and D) provide the highest values for early redemption call options. This is logical given that the inclusion of these on-the-run assets, which are usually traded at lower YTM than the rest due to their high liquidity, in the estimation of the yield curves may imply a downward shift of the same. Lower zero-coupon interest rates on the call options' exercise times imply a greater possibility of exercise and, therefore, a higher call price. In addition, both samples show the highest volatility values for the long term.

This result is especially evident with Sample A (its basket of assets includes all traded assets) for calls with very short terms, those incorporated in 1.25- and 3-year bonds, and with Sample D (its basket of assets only includes 11 on-the-run securities)

[13] The analysis using EGARCH volatilities provide similar results. The corresponding results are available upon request from the authors.

Table 7 Pricing callable bonds

Term to 1st call	Term to 2nd call	Term to maturity	(V$_i$ − Avg)/Avg	A (%)	B (%)	C (%)	D (%)
Noncallable bond		1.25	Bond price	0.00	0.00	− 0.02	0.02
			Bond YTM	− 0.03	0.05	0.22	− 0.24
Callable bond			Call on price ($)	16.02	− 1.15	0.91	4.50
0.25	–	1.25	Call on YTM (bp)	16.02	− 1.15	0.91	4.50
			Delta	27.30	− 2.29	− 2.47	− 3.06
Noncallable bond		3	Bond price	0.04	− 0.03	− 0.04	0.03
			Bond YTM	− 0.62	0.30	0.45	− 0.13
Callable bond			Call on price ($)	17.08	− 9.55	3.42	6.74
0.5	–	3	Call on YTM (bp)	17.03	− 9.52	3.47	6.71
			Delta	17.65	− 0.05	− 0.08	0.13
Noncallable bond		10	Bond price	− 0.06	− 0.17	− 0.15	0.37
			Bond YTM	0.15	0.45	0.36	− 0.96
Callable bond			Call on price ($)	− 0.26	− 0.68	− 0.40	15.64
2	–	10	Call on YTM (bp)	− 0.17	− 0.52	− 0.26	15.24
			Delta	5.48	− 3.88	− 3.85	16.27
Noncallable bond		20	Bond price	− 0.19	− 0.38	− 0.43	1.05
			Bond YTM	0.28	0.57	0.65	− 1.57
Callable bond			Call on price ($)	4.12	− 7.58	− 10.07	29.00
5	6	20	Call on YTM (bp)	4.40	− 7.16	− 9.67	27.83
			Delta	8.19	− 4.64	− 6.33	17.51
Callable bond			Call on price ($)	− 4.59	− 11.00	− 18.62	50.92
18	19	20	Call on YTM (bp)	− 4.33	− 10.47	− 18.12	49.57
			Delta	1.64	− 6.55	− 10.85	31.55
Noncallable bond		30	Bond price	− 0.22	− 0.52	− 0.66	1.48
			Bond YTM	0.25	0.63	0.81	− 1.77
Callable bond			Call on price ($)	5.11	− 12.25	− 12.65	42.66
3	4	30	Call on YTM (bp)	5.39	− 11.57	− 11.88	40.81
			Delta	10.26	− 5.58	− 6.69	23.24
Callable bond			Call on price ($)	5.93	− 11.69	− 12.42	40.84
5	–	30	Call on YTM (bp)	6.20	− 11.02	− 11.67	39.02
			Delta	6.69	− 11.41	− 4.46	31.06
Callable bond			Call on price ($)	8.94	− 11.80	− 12.78	37.63
10	–	30	Call on YTM (bp)	9.15	− 11.25	− 12.14	36.14
			Delta	10.84	− 10.31	− 9.62	30.41

Call prices and call deltas writing in theoretical callable bonds obtained from BDT estimations using zero coupon rates and "historical volatilities" from the four considered datasets: Sample A (full sample), Sample B (full sample excluding on-the-run and first off-the-run), Sample C (mimicking FRB, excluding bills, on-the-run, first off-the-run), and Sample D (mimicking DoT, only on-the-run, 11 maturities). The call price is the difference between the prices of a noncallable and a callable default-free bonds. The callable bond includes one or two at par call options. Both securities are 5% semiannual coupon bonds. The call delta is the ratio between the change in price of the call option and the change in price of the underlying bond. Figures are average relative variations $[(V_{jt} - AvgV_t)/AvgV_t]$ of the value obtained for each sample j on date t (V_{jt}) with respect to the average value of the variable for the four samples on date t $(AvgV_t = \Sigma_j V_{jt})$. Sample period: first working day of March and November from 1996 to 2006 for bonds maturing in 20 and 30 years. For the latter long-term bonds, the differences in call prices obtained with Sample D with respect to the average of all samples is always greater than 29%.

The largest differences in relative terms are obtained for the 20-year bond with calls in years 18 and 19. In this case, the difference between the price of the noncallable bond and the callable bond is on average 50.9% higher in Sample D, which mimics the basket composition of DoT, than its average value for the four yield curves. Sample C, which mimics the basket of assets used by FRB, obtains a price for the call 18.6% below the average. The average call prices embedded in this callable bond obtained from samples D and C are respectively \$1.23 and \$0.64. On the other hand, the largest differences in absolute terms appear for the 30-year bond callable at 10-year. The call option is valued on average at \$5.03 from Sample D and \$3.33 from Sample C. These results show the important economic implications of using one or another sample of data when valuing callable bonds.

4.3 The impact on the forward rates correlation

We also examine the potential consequences of the use of a concrete interest rate dataset on the correlations between pairs of forward rates. Desirable properties of any yield curve are that it should produce a smooth forward rate curve, which converges to a fixed limit as maturity increases. Forward rates are highly sensitive to the shape of the yield curve, particularly at the very long end. Economic rationality suggests an almost perfect positive correlation between long-term forward rates. Under the assumption that forward rates reflect expectations about future short interest rates, agents should perceive similar values for the interest rates for different long horizons. Forward rates play a key role in many financial issues, such as the implementation of interest rate models (for instance, Black or Heath–Jarrow–Morton models), in many product valuations where correlations among forward rates are crucial (e.g., swaptions), and in analyses of market expectations about future short-term interest rates.

Panel A of Table 8 summarizes statistics of pairwise correlations from daily contemporaneous forward rates on various horizons. We compute time series of pairwise correlation coefficients by using 3-month windows between three pairs of several year-ahead forward rates with a 6-month tenor. In particular, we examine correlations between two short-term forward rates (the 6-month spot rate and the 1-year ahead forward rate with a half-year tenor, $\rho(R_{0.5}; F_{1,1.5})$), correlations between two medium-term forward rates, $\rho(F_{2,2.5}; F_{5,5.5})$, and correlations between long-term forward rates, $\rho(F_{10,10.5}; F_{29.5,30})$. As in the previous section, we also report tests of equality mean, median and variance obtained from non-overlapping 3-month periods (Panel B of Table 8).

Panel A shows that at least in levels, values for the four samples are broadly equivalent. These smooth forward rates are likely due to the use of the Svensson method as fitting technique. Independently of the considered set of assets, the weighted version of the Svensson method provides correlation coefficients between forward rates that meet the desirable property to be close to the unity for long maturities that an asymptotically flat forward rate curve should provide. In fact, at the long end average

Table 8 Summary statistics of forward correlations of yield curve estimates using Svensson (1994)

t	Mean				Std.dev.				Median			
	A	B	C	D	A	B	C	D	A	B	C	D
Panel A: Descriptive statistics												
$R_{0.5}$; $F_{1,1.5}$	0.51	0.49	0.61	0.37	0.22	0.22	0.28	0.32	0.53	0.53	0.72	0.46
$F_{2,2.5}$; $F_{5,5.5}$	0.77	0.78	0.88	0.48	0.15	0.15	0.08	0.34	0.83	0.84	0.91	0.53
$F_{10,10.5}$; $F_{29.5,30}$	0.73	0.85	0.83	0.79	0.28	0.07	0.18	0.16	0.83	0.86	0.89	0.83

t	Skewness				Excess kurtosis			
	A	B	C	D	A	B	C	D
Panel A: Descriptive statistics								
$R_{0.5}$; $F_{1,1.5}$	− 0.38	− 0.45	− 1.93	− 1.11	− 0.68	− 0.53	3.60	1.10
$F_{2,2.5}$; $F_{5,5.5}$	− 1.07	− 1.11	− 2.19	− 0.50	0.19	0.11	5.39	− 0.52
$F_{10,10.5}$; $F_{29.5,30}$	− 1.80	− 1.04	− 1.93	− 2.54	2.23	2.18	2.63	7.52

t	Ratio A/C						Ratio B/C					
	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test
Panel B: Test results for differences in forward correlations												
$R_{0.5}$; $F_{1,1.5}$	0.76	31	− 2.01*	2.45**	3.41***	0.40	0.83	27	− 1.93*	3.03***	3.62***	0.22
$F_{2,2.5}$; $F_{5,5.5}$	0.93	15	− 4.83***	4.76***	4.63***	23.01***	0.94	15	− 5.07***	4.76***	4.82***	25.55***

Table 8 continued

t	Ratio A/C						Ratio B/C					
	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test
$F_{10,10.5}$; $F_{29.5,30}$	0.93	25	− 0.25	3.32***	2.73***	3.26*	0.96	25	1.55	3.32***	2.53***	9.42***

t	Ratio C/D					
	Mean	% > 1.0	*t*-test	Sign test	Signed-rank test	Variance test
Panel B: Test results for differences in forward correlations						
$R_{0.5}$; $F_{1,1.5}$	1.28	73	2.33**	3.03***	2.86***	1.61
$F_{2,2.5}$; $F_{5,5.5}$	1.55	92	2.29**	5.63***	4.64***	87.75***
$F_{10,10.5}$; $F_{29.5,30}$	1.08	79	1.69*	3.90***	2.70***	0.50

Pairwise correlation coefficients from daily forward rates using 30-day rolling windows. They are 6-month forward rate starting at time 0, 1, 2, 5, 10 and 29.5 ($F_{0,0.5} = R_{0.5}$, $F_{1,1.5}$, $F_{2,2.5}$, $F_{5,5.5}$, $F_{10,10.5}$, and $F_{29.5,30}$ respectively). For instance, $F_{5,5.5}$ is the 5-year ahead forward rate with a half-year tenor. As the input, we use zero-coupon interest rates from our own yield curve estimates, applying the weighted Svensson to the four different datasets from GovPx asset prices (Sample A: full sample (bills, notes, and bonds); Sample B: full sample excluding on-the-run and first off-the-run; Sample C: mimicking FRB, excluding bills, on-the-run, first off-the-run; Sample D: mimicking DoT, only on-the-run, 11 maturities). We consider 2864 daily estimates of the term structure of interest rates for each basket of assets during the period between January 1996 and December 2006

Test for differences in forward correlations are computed using 3-month non-overlapping periods. The X/Y ratio is defined as the mean correlation coefficient for the X dataset divided by the corresponding mean for the Y dataset. The hypothesis that the mean of the X/Y ratio is equal to unity is tested using a standard *t*-test. A standard sign test is used to test that the proportion of X/Y ratios higher than unity is equal to the proportion of ratios lower than unity, which is equal to 0.5. The Levene statistic is used to test the equality of variances between X and Y. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively, in a two-tailed test correlations are always greater than 0.73. We can also see that all the correlations are always positive, and their distribution has negative skewness. The excess kurtosis denotes that correlation between the longest forward rates provides extreme deviations. In the sample without bills, on-the-run, and first off-the-run bonds (Sample C), the problem of fat tail on the left side also appears for all the correlations.

Only Sample D seems to produce clearly lower correlations at the short end of the yield curve, a result that can be caused by the more accurate adjustment of the zero-coupon rates that this sample provides in this tranche of the yield curve. Finally, tests of equality in mean, median and variance (Panel B) show significant differences in all the analyzed pairs of samples and maturities. Thus, the apparently similar descriptive values among the four samples are significantly different, which could have important implications for pricing complex interest derivatives.

4.4 The impact on a simple test of the expectation hypothesis

As another illustration of the relevance of using a particular yield curve estimate in this section, we replicate the test of one of the most relevant hypotheses about the information content of the term structure of interest rates: the expectation hypothesis.

Empirical researchers have frequently rejected the expectations hypothesis. However, results vary from one study to the next depending on the methodology, the maturity spectrum, and the sample period. For instance, Campbell and Shiller (1991) reject the expectations hypothesis for all combinations of short- and long-term rates when the maturity of the long-term rate is less than 4 years. However, they only reject the hypothesis in one case when the long-term maturity is 4 years or larger. Examining the very short end of the term structure, Longstaff (2000) finds support for the expectations hypothesis from repo data. This author observes that when using short-term yields data that do not suffer from the specialness/idiosyncracy of the Treasury markets, the expectation hypothesis, unlike the case of Treasury yields, cannot be rejected. However, Downing and Oliner (2007) and Brown et al. (2008) obtain the opposite result from commercial paper interest rates.

The interest in and relevance of this issue among academics and practitioners is obvious. In this section, we only attempt to analyze the impact of using one or another yield curve dataset on the result of a simple expectations hypothesis test. This study does not claim to provide a comprehensive and detailed analysis of the expectations analysis but instead aims to highlight the implications of using different samples of rates. Therefore, we merely replicate the classical Campbell's (1995) test for our alternative yield curve datasets. As a slight robustness improvement, we use an instrumental variable regression. Campbell and Shiller (1991) suggest this technique because otherwise, the regression results should be extremely sensitive to measurement error in the long-term interest rate.

We explore the expectation hypothesis by using the same methodology, the same interest rate maturities, and the same sample period, but using our four alternative interest rate datasets. We do not attempt to find evidence supporting or rejecting the theory; instead, we only attempt to illustrate the possible implications of considering one or another yield curve dataset.

Following the detailed description of Campbell (1995), we compute continuously compounded yields by using 1 month as the basic time unit. The yield spread of a bond is defined as the difference between its yield and the short rate. As the short rate, Campbell (1995) uses the yield of a 1-month Treasury bill. Because the DoT reports yields of on-the run securities for several fixed maturities, our actual 1-month yield is proxied by the 1-month yield reported by the DoT dataset.[14] The excess return is calculated as the difference between bond returns and the short yield. The excess return is also the yield spread less $(m - 1)$ times the change in the bond yield, where m is the maturity in months.

It is well known that the expectations hypothesis implies that excess returns of long bonds over short bonds are unforecastable, with a zero mean in the case of the pure expectations hypothesis. Panel A of Table 9 reproduces Table 1 (p. 135) in Campbell (1995). The first row checks whether the excess return of long bonds over short bonds has a zero mean.[15] We replicate this table by using our Samples A to D to isolate the impact of the different basket of assets; the outcomes are shown in Panels A, B, C and D of Table 10. The excess returns in Campbell's original results show a hump with a maximum at 12-month maturity (Panel A, Table 9). Instead of a hump, we observe that the excess return increases with maturity. This outcomes show important differences among datasets in the estimation of excess returns for maturities shorter than or equal to 12 months. The shorter the maturity, the greater the difference. These discrepancies are especially acute in the case of Sample C, the one that tries to mimic FRB. Removing bills from the basket of assets produces important differences with respect to other datasets.

Campbell (1995) also analyses whether the yield spread has forecasting power with respect to future short interest rates. Panel B of Table 9 reproduces Campbell's (1995) original Table 2 (p. 139) with the results of two alternative tests of the expectations hypothesis. Panels A to D of Table 11 replicate these tests using our four datasets. Excluding bills (Sample C) produces significant differences among the results obtained from the rest of yield curves in all maturities in both level and statistical significance. Tests using the sample of only on-the-run securities (Sample D) also provide different values from those obtained using the other yield curves. Divergences affect especially short maturities.

In summary, the results obtained in this section suggest that the conclusions of a simple test as those conducted by Campbell (1995) can differ significantly depending upon the yield curve estimates used to perform the test. The differences observed may affect both the size of the parameters estimations and its statistical significance, especially at the short end of the yield curve. These results parallel to those of Andreasen

[14] The DoT provides market yields at fixed maturities calculated from composites of quotations obtained by the Federal Reserve Bank of New York. According to the information on the DoT website, the yields are either investment yields or bond-equivalent yields. The formula uses simple interest and day count convention actual/actual. We recalculate this yield as a continuously compound interest rate. No 1-month rates are available in this dataset until August 1, 2001. Prior to this date, we estimate the corresponding interest rate by using cubic interpolation.

[15] The data in Table 7 are reported in annualized percentage points, i.e., the natural monthly variables are multiplied by 1200.

Table 9 Original tables reported by Campbell (1995) (p. 135, 139)

Variable	Long bond maturities (months)						
	2	3	6	12	24	48	120
Panel A: Original Campbell's "Table 1 means and standard deviations of term structure variables" (Campbell 1995, p. 135)[a]							
Excess return	0.379	0.553	0.829	0.862	0.621	0.475	− 0.234
	(0.640)	(1.219)	(2.950)	(6.203)	(11.29)	(19.32)	(36.77)
Change in yield	0.014	0.014	0.014	0.014	0.014	0.014	0.013
	(0.591)	(0.575)	(0.569)	(0.546)	(0.486)	(0.408)	(0.307)
Yield spread	0.196	0.324	0.569	0.761	0.948	1.141	1.358
	(0.210)	(0.301)	(0.437)	(0.594)	(0.799)	(1.013)	(1.234)
Panel B: Original Campbell's "Table 2 regression coefficients" (Campbell 1995, p. 139)[b]							
Short-run changes in long yields	0.019	− 0.135	− 0.842	− 1.443	− 1.432	− 2.222	− 4.102
	(0.194)	(0.285)	(0.444)	(0.598)	(0.996)	(1.451)	(2.083)
Long-run changes in short yields	0.510	0.473	0.301	0.253	0.341	0.435	1.311
	(0.097)	(0.149)	(0.147)	(0.210)	(0.221)	(0.398)	(0.120)

[a]Campbell footnote: "*Source*: Author's calculations using estimated monthly zero-coupon yields, 1952–1991, from McCulloch and Kwon (1993). The data are measured monthly, but expressed in annualized percentage points. Each row shows the mean of the variable, with the standard deviation below in parentheses. Excess returns and yield spreads are measured relative to 1-month Treasury bill rates."

[b]Campbell footnote: "*Source*: Author's calculations using estimated monthly zero-coupon yields, 1952–1991, from McCulloch and Kwon (1993). Each row shows a regression coefficient β, with the standard error below in parentheses. Each coefficient should be one if the expectations hypothesis holds. The regression in the first row is $y_{m-1,t+1} - y_{m,t} = \alpha + \beta\left(y_{m,t} - y_{1,t}\right)/(m-1)$ where m is long bond maturity in months. The regression in the second row is $\sum_{i=1}^{m-1} \frac{y_{1,t+i}}{m-1} - y_{1,t} = \alpha + \beta\left(\frac{m-1}{m}\right)\left(y_{m,t} - y_{1,t}\right)$. The standard error in the second row is corrected for serial correlation in the error term of the regression."

et al. (2017) in their comparison of dynamic term structure models estimated using two approaches to obtain the yield curve.

5 Conclusions

When comparing alternative datasets that contain information about zero-coupon rates, researchers usually focus on the ability to replicate bond prices or to adjust dynamic term structure models. However, little attention has been paid to the impact of dataset choice on other variables that are particularly relevant to finance practitioners and academics such as the resulting volatilities and correlations among different rates that are implicated in these datasets.

There are primarily two potential sources of discrepancies among them. On the one hand, some datasets are obtained using different models to describe the discount function, including both parametric and non-parametric models. On the other hand, the information used to estimate zero-coupon rates differ from one dataset to other in at least two aspects. First, the bond prices used can be quotes, mid prices or even

Table 10 Replication of "Table 1 means and standard deviations of term structure variables" (Campbell 1995, p. 135)

Variable	Long bond maturities (months)						
	2	3	6	12	24	48	120
Panel A: Sample A (Svensson model from GovPx bond dataset: full sample of bills, notes, and bonds)							
Excess return	– 0.022	0.041	0.196	0.416	0.731*	1.260*	2.606
	(0.291)	(0.439)	(0.955)	(2.479)	(6.368)	(14.42)	(31.86)
Change in yield	– 0.003	– 0.002	– 0.001	– 0.001	– 0.003*	– 0.005*	– 0.007
	(0.249)	(0.223)	(0.201)	(0.228)	(0.276)	(0.305)	(0.267)
Yield spread	– 0.025	0.037	0.190	0.401	0.651*	0.999*	1.718
	(0.198)	(0.224)	(0.306)	(0.413)	(0.552)	(0.807)	(1.273)
Panel B: Sample B (Svensson model from Sample A, excluding on-the-run and first off-the-run)							
Excess return	– 0.025	0.040	0.199	0.420	0.734*	1.264*	2.618
	(0.291)	(0.439)	(0.956)	(2.484)	(6.373)	(14.41)	(31.85)
Change in yield	– 0.003	– 0.002	– 0.001	– 0.001	– 0.003*	– 0.005*	– 0.007
	(0.250)	(0.224)	(0.202)	(0.229)	(0.276)	(0.305)	(0.267)
Yield spread	– 0.027	0.036	0.193	0.405	0.654*	1.002*	1.726*
	(0.196)	(0.224)	(0.308)	(0.416)	(0.551)	(0.808)	(1.272)
Panel C: Sample C (Svensson model mimicking FRB basket of assets, i.e., sample excluding bills, on-the-run and first off-the-run)							
Excess return	0.133	0.169	0.269	0.440	0.732*	1.261*	2.631
	(0.526)	(0.666)	(1.190)	(2.547)	(6.341)	(14.38)	(31.99)
Change in yield	– 0.000	– 0.000	– 0.001	– 0.002	– 0.003*	– 0.005*	– 0.008
	(0.279)	(0.263)	(0.235)	(0.234)	(0.275)	(0.304)	(0.268)
Yield spread	0.134	0.170	0.265	0.421	0.653*	1.005*	1.719*
	(0.444)	(0.432)	(0.415)	(0.436)	(0.550)	(0.804)	(1.271)
Panel D: Sample D (Svensson model mimicking DoT basket of assets, i.e., on-the-run bills and bonds for 11 maturities)							
Excess return	– 0.071*	0.017	0.191	0.406	0.714*	1.197*	2.586
	(0.379)	(0.528)	(1.047)	(2.554)	(6.306)	(14.92)	(31.37
Change in yield	– 0.002	– 0.001	– 0.000	– 0.001	– 0.004*	– 0.004*	– 0.008*
	(0.314)	(0.259)	(0.214)	(0.233)	(0.274)	(0.316)	(0.263)
Yield spread	– 0.073	0.014	0.190	0.390	0.627*	0.977*	1.652*
	(0.205)	(0.230)	(0.323)	(0.424)	(0.559)	(0.814)	(1.254)

Continuously compounded yields calculated from zero-coupon interest rates from our own yield curve estimates, applying the weighted Svensson to the four different datasets from GovPx asset prices. The sample period ranges from January 1996 to December 2006. The yield spread is defined as the difference between the bond yield and the short yield. The excess return is calculated as the difference between the return on a bond and the short yield. The short yield is proxied by the 1-month yield reported by the DoT dataset

*Indicates significance at 5% level in a two-tailed test

Table 11 Replication of "Table 2 regression coefficients" (Campbell 1995, p. 139)

Variable	Long bond maturities (months)						
	2	3	6	12	24	48	120
Panel A: Sample A (Svensson model from GovPx bond dataset: full sample of bills, notes, and bonds)							
Short-run changes in long yields	0.872*	1.334*	1.776*	1.694*	0.229	– 2.036	– 2.797
	(0.352)	(0.396)	(0.476)	(0.788)	(1.327)	(1.811)	(2.326)
Long-run changes in short yields	2.398*	1.874*	1.438*	1.376*	1.131*	0.952*	– 0.074
	(0.742)	(0.372)	(0.204)	(0.216)	(0.273)	(0.202)	(0.143)
Panel B: Sample B (Svensson model from Sample A excluding on-the-run and first off-the-run)							
Short-run changes in long yields	0.880*	1.296*	1.659*	1.594*	0.224	– 2.055	– 2.826
	(0.358)	(0.387)	(0.460)	(0.776)	(1.330)	(1.809)	(2.329)
Long-run changes in short yields	2.447*	1.855*	1.383*	1.328*	1.128*	0.952*	– 0.074
	(0.753)	(0.367)	(0.201)	(0.215)	(0.274)	(0.202)	(0.142)
Panel C: Sample C (Svensson model mimicking FRB basket of assets, i.e., sample excluding bills, on-the-run and first off-the-run)							
Short-run changes in long yields	0.090	0.224	0.748*	1.309	0.082	– 2.002	– 2.687
	(0.078)	(0.152)	(0.350)	(0.721)	(1.328)	(1.812)	(– 1.148)
Long-run changes in short yields	0.585*	0.577*	0.810*	1.175*	1.120*	0.964*	– 0.079
	(0.171)	(0.136)	(0.140)	(0.201)	(0.276)	(0.203)	(0.145)
Panel D: Sample D (Svensson model mimicking DoT basket of assets, i.e., on-the-run bills and bonds for 11 maturities)							
Short-run changes in long yields	1.009*	1.531*	1.866*	1.722*	0.394	– 2.048	– 2.651
	(0.467)	(0.487)	(0.565)	(0.848)	(1.279)	(1.852)	(2.329)
Long-run changes in short yields	2.154*	1.957*	1.425*	1.461*	1.125*	0.913*	– 0.080
	(0.792)	(0.387)	(0.213)	(0.220)	(0.261)	(0.200)	(0.143)

Continuously compounded yields calculated from zero-coupon interest rates from our own yield curve estimates, applying the weighted Svensson to the four different datasets from GovPx asset prices. The sample period ranges from January 1996 to December 2006. The yield spread is defined as the difference between the bond yield and the short yield. The excess return is calculated as the difference between the return on a bond and the short yield. The short yield is proxied by the 1-month yield reported by the DoT dataset

*Indicates significance at 5% level in a two-tailed test

yields to maturity. Second, the basket of assets used to estimate yield curves also varies considerably.

In this paper, we examine whether using different inputs to estimate zero-coupon rates has consequences in two relevant variables: the resulting volatility term structure and the correlations between pairs of forward rates. DoT, FRB and F082 interest rate datasets use different fitting techniques, but also use really different sets of assets. DoT considers only on-the-run securities, including four maturities of the most recent auction bills. FRB excludes bills and the on-the-run and first-off-the-run bonds. F082 includes all the outstanding bonds, even old callable bonds.

We isolate the impact of the set of assets used in the yield curve construction as source of discrepancies between these popular datasets. To perform this analysis we estimate zero-coupon yields using the same methodology and bond prices but attempt to mimic the basket of assets employed by the three initial zero-coupon datasets. Depending on the initial data choice, we are actually estimating different interest rates: the spot rates corresponding to the average market liquidity level, the spot rates of the most liquid references, and the spot rates of the seasoned bonds. Even the inclusion or exclusion of the bills has also relevant implications.

Our analysis includes the analysis of the behavior of the parameters estimated in the Svensson's (1994) model fitting, the study of the pricing errors of assets with different degrees of liquidity and covering the whole spectrum of terms to maturity, the in-depth analysis of the behavior of the volatility term structure of the zero coupon rates generated, the pricing of different callable bonds and the analysis of the correlations between forward rates.

We obtain evidence that suggests that including all bills and bonds (Sample A) provides the best fit for the short end of the yield curve. Removing the on-the-run bonds (Sample B) only improves the in-sample performance of the fitting model for medium maturities but deteriorates in fit long maturities. If the bills are also deleted (Sample C, which reproduces the asset composition of the FRB database), the adjustment at both ends of the yield curve worsens. This sample provides the best fit for intermediate maturities and causes the resulting volatility of all interest rates to decrease. However, it provides extreme deviations in the correlation between the longest forward rates. The estimation from only on-the-run bills and bonds (Sample D, which mimics the composition of the DoT database) performs well for long maturities and produces much higher estimates of volatility, higher prices of interest rate derivatives and lower forward rate correlations than any other sample at both ends of the yield curve.

The outcomes suggest that even if bond prices are replicated accurately using any of the datasets, the impact on other variables can be remarkable, especially at both ends of the yield curve. It is worth noting the large effects of the basket of assets on the volatility term structure, regardless of whether it is estimated from a simple standard deviation or from EGARCH models. Important differences may arise in this and other key financial variables depending on the securities included in the yield curve estimations.

Finally, to test the relevance of this paper's findings, we price callable bonds and replicate some well-known tests of the expectations hypothesis using our own estimates of the yield curve using different baskets of Treasury assets. The outcomes clearly indicate that results may differ significantly depending upon the set of yield curves used. The main discrepancies appear in the pricing of long-term callable bonds and when short rates (less than 1 year) are involved to test the expectations hypothesis.

In conclusion, we can state that the particular set of securities employed to estimate the term structure of zero-coupon rates can lead to dynamics and a relationship among interest rates that can differ significantly. The impact of choosing one or another dataset is particularly notable at both the short and the long end of the yield curve. Although users of yield curves (traders, portfolio managers, researchers, policymakers, etc.) are all aware of the fact that constructed zero-coupon yields are "just" representations of the underlying yield curve, we obtain evidence that the use of one or another yield curve data set may lead to different results on pricing, testing hypotheses or models, and managing risk practices. The choice of the yield curve data set may become a nontrivial decision. Results of our concrete analysis could suggest that Sample D, which is fitted with the least computational effort from only eleven on-the-run bills and bonds used by the DoT, is probably the best alternative for many financial proposals.

Acknowledgements We are indebted to Zvika Afik, Alain Coën, Manfred Frühwirth, Alois Geyer, Robert Korajczyk, John Merrick, Pietro Millosovich, Alfonso Novales, Christian Speck, Christian Wagner, Josef Zechner and Hairui Zhang for very detailed suggestions. In addition, we would like to thank the anonymous referee for their constructive and valuable comments on this paper. This work was supported by the Spanish *Ministerio de Economía, Industria y Competitividad* (ECO2017-89715-P). Any errors are solely the responsibility of the authors.

Appendix: Technical details of our dataset of government securities prices

GovPX consolidates and posts real-time quotes and trades data from six of the seven major interdealer brokers (with the notable exception of Cantor Fitzgerald).[16] Collectively, these brokers account for approximately two-thirds of the voice interdealer broker market. In turn, the interdealer market is approximately one-half of the total market (see Fleming 2003). Thus, although the estimated bills coverage exceeds 90% in every year of the Fleming's GovPX sample (Jan 97–Mar 00), the availability of 30-year bond data is limited because of Cantor Fitzgerald's prominence in the long-maturity segment of the market. According to Mizrach and Neely (2006), voice-brokered trading volume began to decline after 1999 as electronic trading platforms (e.g., eSpeed, BrokerTec) became available. Indeed, GovPX does not provide aggregate volume and transaction information after May 2001. After ICAP's purchase of GovPX in January 2005, ICAP PLC was the only broker reporting through GovPX. Therefore, we assume an imperceptible impact from the decline in GovPX market coverage on our estimates because we consider the midpoint prices and yields between the bid and ask prices at 5 pm.

The GovPX dataset contains snapshots of the market situation at 1 pm, 2 pm, 3 pm, 4 pm, and 5 pm. Each snapshot includes detailed individual security information, such as CUSIP, coupon, maturity date, and product type (an indicator of whether the security is trading when issued, on the run, or active off the run). The transaction data available until May 2001 include the last trade time, size, and side (buy or sell), the price (or yield in the case of bills), and the aggregate volume (volume in millions traded from 6 pm on the previous day to 5 pm). The quote data used from June 2001 include the best bid and ask prices (or discount rate actual/360 in the case of bills) and the mid-price and mid-yield (actual/365).

Our initial sample relies on the information at 5 pm, i.e., the last transaction during "regular trading hours" (from 7:30 am to 5:00 pm Eastern Time, ET), if available, and quote data otherwise. Quote prices are used during the last part of the sample period because trading volume information is not reported by GovPX. We complement the GovPX data with official data on the dates for the last issue and the first coupon

[16] GovPX Inc. was established under the guidance of the Public Securities Association as a joint venture among voice brokers in 1991 to increase public access to US Treasury security prices.

payment and the coupon rate of each Treasury security.[17] This information is publicly available on the US Treasury website.

To improve the adjustment at the short end of the yield curve, we initially consider all Treasury bills. In this maturity segment, bills are very more actively traded than old off-the-run notes and bonds, most of which are absorbed into investors' inactive portfolios.[18] Liquidity differences in these short-term assets imply large divergences in yields to maturity. Thus, we include only Treasury notes and bonds that have at least 1 year of life remaining. However, we are forced to modify this criterion from 2002 because the number of outstanding bills with terms to maturity between 6 months and 1 year declines considerable during the year 2000, and the 1-year Treasury bill is no longer auctioned as of the beginning of March 2001. Therefore, we also consider Treasury notes and bonds with remaining maturities between 6 and 12 months for the period after 2001.

We also apply other data filters that are designed to enhance the quality of the data. First, we do not include transactions that are associated with "when-issued" and cash management or trades and quotes that are related to callable and flower bonds and TIPS. Second, when two or more securities have the same maturity, we consider only trades and quotes for the youngest security, i.e., the security with the last auction date. Finally, we exclude yields that differ greatly from yields at nearby maturities. We apply an ad hoc filter based on common sense.[19] In addition, we occasionally observe that deleting a single data point in the set of prices used to fit the yield curve can produce a notable shift in both the parameters and the fitted yields, notably improving the fit. This phenomenon is also noted by Anderson and Sleath (1999).

To control for market conventions, we recalculate the price of each security in a homogeneous procedure to prevent a situation in which the effects of different market conventions depend on maturities and assets. Every price is valued at the trading date on an actual/actual day-count basis. In the case of Treasury bills, we first obtain the price at the settlement date from the last trade price, if available, and from the mid-price between the bid and the ask prices otherwise.[20] In both cases, the GovPX reported price is a discount rate using the actual/360 basis. Second, we compute the yield-to-maturity as a compound interest rate by using the actual/actual. Third, we calculate "our" price at the trading date by using the yield-to-maturity obtained in the previous step.[21] In the case of Treasury notes and bonds, the price is directly reported in the

[17] The "standard interest payment" field indirectly provides information to identify callable bonds and TIPS (Treasury Inflation-Protected Securities). We exclude these assets.

[18] See, for instance, Sarig and Warga (1989) and Kamara (1994). In addition, Fleming (2003) emphasizes that GovPX bill coverage is larger than bond and note coverage.

[19] To fix a universal criterion for the entire maturity spectrum and sample period, such as excluding trades for which price or yield-to-maturity (YTM) exceeds certain percentage of the neighbour trades, does not work. Numerous factors imply differences in prices and YTM. For instance, assets with the same maturity and liquidity level but different coupon amounts are traded at different prices; turmoil and flight-to-liquidity periods involve high volatile prices; on-the-run bonds are often traded with a large liquidity premium and should not be considered outliers because they are the most actively traded assets and the market benchmark, etc. Additionally, there are interdealer brokers' posting errors similar to those mentioned by Fleming (2003).

[20] We do not consider the reported mid yield, which is simple interest with an actual/365 basis, except for more than 6-month remaining maturity bills, which are valued using the bond equivalent yield.

[21] Note that the settlement date is generally one working day after the trading date.

data as the last trade price or the mid-price. From this price, we apply the mentioned second and third steps to obtain "our" homogeneous price.[22]

References

Amihud, Y., & Medelson, H. (1991). Liquidity, maturity and the yield on U.S. Treasury securities. *Journal of Finance, 46,* 1411–1425.

Andersen, T. G., & Benzoni, L. (2007). Do bonds span volatility risk in the U.S. Treasury market? A specification test for affine term structure models. NBER Working Paper No. W12962.

Anderson, N., & Sleath, J. (1999). New Estimates of the U.K. Real and Nominal Yield Curves. *Bank of England Quarterly Bulletin*, November, 384–392.

Andreasen, M. M., Christensen, J. H. E., & Rudebusch, G. D. (2017). Term structure analysis with big data. Federal Reserve Bank of San Francisco, Working Paper 2017-21.

Bank of International Settlements. (2005). Zero-coupon yield curves: technical documentation. BIS Papers, No. 25, October.

Barnejee, S., & Graveline, J. J. (2013). The cost of short-selling liquid securities. *The Journal of Finance, 68,* 637–664.

Black, F., Dermand, E., & Toy, W. (1990). A one-factor model of interest rates and its application to treasury bond options. *Financial Analysts Journal, 46,* 33–39.

Bliss, R. R. (1996). Testing term structure estimation methods. *Advances in Futures and Options Research, 9,* 197–231.

Bolder, D. J., & Gusba, S. (2002). Exponentials, polynomials, and Fourier series: More yield curve modelling at the bank of Canada. Bank of Canada, Working Paper 2002-29.

Bolder, D. J., Johnson, G., & Metzler, A. (2004). An empirical analysis of the Canadian term structure of zero-coupon interest rates. Bank of Canada Working Paper No. 2004-48.

Brown, C. G., Cyree, K. B., Griffiths, M. D., & Winters, D. B. (2008). Further analysis of the expectations hypothesis using very short-term rates. *Journal of Banking & Finance, 32,* 600–613.

Campbell, J. Y. (1995). Some lessons from the yield curve. *Journal of Economic Perspectives, 9*(3), 129–152.

Campbell, J., & Shiller, R. (1991). Yield spreads and interest rates movements: A bird's eye view. *Review of Economic Studies, 58,* 495–514.

Cherian, J. A., Jacquier, E., & Jarrow, R. A. (2004). A model of the convenience yields in on-the-run treasuries. *Review of Derivatives Research, 7,* 29–97.

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews, 27*(1–3), 46–78.

Díaz, A., & Escribano, A. (2017). Liquidity measures throughout the lifetime of the U.S. Treasury bond. *Journal of Financial Markets, 33,* 42–74.

Díaz, A., Jareño, F., & Navarro, E. (2011). Term structure of volatilities and yield curve estimation methodology. *Quantitative Finance, 11*(4), 573–586.

Díaz, A., Merrick, J. J., & Navarro, E. (2006). Spanish Treasury bond market liquidity and volatility pre- and post-European Monetary Union. *Journal of Banking & Finance, 30,* 1309–1332.

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics, 130*(2), 337–364.

Downing, C., & Oliner, S. (2007). The term structure of commercial paper rates. *Journal of Financial Economics, 83,* 59–86.

Duffie, D. (1996). Special repo rates. *The Journal of Finance, 51*(2), 493–526.

Fleming, M. J. (2000). Financial market implications of the Federal debt paydown. Staff Report, Federal Reserve Bank of New York, No. 120.

Fleming, M. J. (2003). Measuring Treasury market liquidity. *Federal Reserve Bank of New York Economic Policy Review, 9,* 83–108.

Fontaine, J. S., & García, R. (2012). Bond liquidity premia. *Review of Financial Studies, 25,* 1207–1254.

Goldreich, D., Hanke, B., & Nath, P. (2005). The price of future liquidity: Time-varying liquidity in the U.S. Treasury market. *Review of Finance, 9,* 1–32.

[22] We control for the special size of the first interest payment in just-issued securities.

Graveline, J. J., & McBrady, M. R. (2011). Who makes on-the-run Treasuries special? *Journal of Financial Intermediation, 20,* 620–632.

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

Jordan, B., & Jordan, S. (1997). An empirical investigation of the repo special rate. *The Journal of Finance, 52,* 2051–2072.

Jordan, J. V., & Mansi, S. A. (2003). Term structure estimation from on-the-run Treasuries. *Journal of Banking & Finance, 27*(8), 1487–1509.

Kamara, A. (1994). Liquidity, taxes, and short-term yields. *Journal of Financial and Quantitative Analysis, 29*(3), 403–417.

Krishnamurthy, A. (2002). The bond/old-bond spread. *Journal of Financial Economics, 66*(2), 463–506.

Krishnamurthy, A., & Vissing-Jorgensen, A. (2012). The aggregate demand for Treasury debt. *Journal of Political Economy, 120*(2), 233–267.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics: Essays in honor of Harold Hotelling* (pp. 278–292). Stanford, CA: Stanford University Press.

Longstaff, F. A. (2000). The term structure of very short-term rates: New evidence for the expectations hypothesis. *Journal of Financial Economics, 58*(3), 397–415.

Longstaff, F. A. (2004). The flight-to-liquidity premium in U.S. Treasury bond prices. *Journal of Business, 77,* 511–526.

McCulloch, J., & Kwon, H. (1993). U.S. term structure data, 1947–1991. Ohio State University Working Paper, pp. 93–96.

Mizrach, B., & Neely, C. (2006). The transition to electronic communications networks in the secondary Treasury market. *Federal Reserve Bank of St. Louis Review, 88*(6), 527–541.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*(2), 347–370.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of Business, 60,* 473–489.

Pasquariello, P., & Vega, C. (2009). The on-the-run liquidity phenomenon. *Journal of Financial Economics, 92,* 1–24.

Sack, B. P., & Elsasser, R. (2004). Treasury inflation-indexed debt: A review of the U.S. experience. *FRBNY Economic Policy Review, 10*(1), 47–63.

Sarig, O., & Warga, A. (1989). Bond price data and bond market liquidity. *Journal of Financial and Quantitative Analysis, 24*(3), 367–378.

Svensson, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992–1994. Centre for Economic Policy Research, Discussion Paper 1051.

Vayanos, D., & Weill, P. (2008). A search-based theory of the on-the-run phenomenon. *The Journal of Finance, 63,* 1361–1398.

Warga, A. (1992). Bond returns, liquidity, and missing data. *Journal of Financial and Quantitative Analysis, 27*(4), 605–617.

Yallup, P. J. (2012). Models of the yield curve and the curvature of the implied forward rate function. *Journal of Banking & Finance, 36*(1), 121–135.