Estimating Yield Curve Noise
by
Michael Abrahams

Submitted to the Sloan School of Management
on June 1, 2018, in partial fulfillment of the
requirements for the degree of
MASTER OF SCIENCE IN MANAGEMENT RESEARCH

Abstract
In this paper, I explore methods for estimating noise in the yield curve. I evaluate optimization methods for fitting yield curves using the Nelson-Siegel model where recommendations in the literature remain unclear. I provide open source code on Github including contributions to the QuantLib C++ financial library.

Thesis Supervisor: Adrien Verdelhan
Title: Class of 1956 Career Development Professor
Associate Professor of Finance
## Contents

**Cover page** 1

**Abstract** 3

### Estimating Yield Curve Noise 7

1 Introduction ................................. 7

2 Model .................................. 8

2.1 Nelson-Siegel-Svensson parameterization ............................. 8

2.2 Objective Function ......................... 9

2.3 Noise Measures .......................... 10

3 Estimation ................................ 11

3.1 Data .................................... 11

3.2 Parameter restrictions ....................... 12

3.3 Optimization techniques ..................... 13

3.4 A comment on the visual accuracy of par yield curves ............ 15

4 Results .................................. 16

4.1 Estimated yield curves ...................... 16

4.2 Noise Measures .......................... 18

4.3 Liquidity measures ......................... 21

4.4 Pricing Currency Portfolios ................. 23

5 Conclusion ................................ 24

**Bibliography** 26
Estimating Yield Curve Noise

1 Introduction

On a typical day, several hundred individual Treasury securities are traded with various issue dates, coupon rates, and trading volumes, yet the end-of-day prices quoted for these securities can be modeled almost entirely by simple and smooth parameterized functions known as Nelson-Siegel-Svensson model. This model is typically able to fit the entire cross-section with a standard error of only two to three basis points. However, in a market with trading volumes averaging over 500 billion dollars per day, those final few basis points matter, and the residual yield curve noise contains significant information about broader market conditions. Producing tight estimates of the Treasury yield curve at this level of accuracy requires on a number of detailed modeling decisions. In this paper I explore the effects of different specifications on these yield curve estimates. I explain my experimentation with the conflicting results in the literature concerning optimization methods for the Nelson-Siegel model. To help others take on this task this I am releasing an open source codebase on Github including modifications to the QuantLib C++ financial library, found on Github at [https://github.com/miabrahams/YieldCurveNoise](https://github.com/miabrahams/YieldCurveNoise). Time constraints for this thesis leave numerous questions to further this research.

My paper is closely related to a number of others in the literature. Most prominently is Hu, Pan & Wang (2013), who first documented the information in the yield curve noise and whose work I am extending. The second closely related strand of literature documents the challenge of fitting the Nelson-Siegel-Svensson model to a high degree of accuracy - this literature includes Gilli, Grosse & Schumann (2010), and Manousopoulos and Michalopoulos (2009), who evaluate the use of global optimization techniques, and Annaert et al. (2013),
who propose regularization. The yield-curve noise measure is one of several liquidity proxies that can be extracted from Treasury market data; some of these papers include Fleming (2003), Adrian, Fleming and Vogt (2017), Fleming, Mizrach & Nguyen (2017). The most dramatic movement in relative Treasury pricing occurred during the 2008 financial crisis; the microstructure of this period is documented in detail by Musto, Nini, and Schwarz (2016).

2 Model

2.1 Nelson-Siegel-Svensson parameterization

The Nelson-Siegel-Svensson model (NSS) is a widely used parameterization for reduced form estimation of yield curve data introduced by Svensson (1994). Its popularity derives from its ability to closely match observed yield curves, and a parsimonious specification of only six parameters, and despite this flexibility the parameters can be given a straightforward interpretation. The instantaneous forward rate of maturity $n$ is assumed to take the following form:

$$f(n; \phi) = \beta_0 + \beta_1 e^{-\frac{n}{\tau_1}} + \beta_2 \frac{n}{\tau_1} e^{-\frac{n}{\tau_1}} + \beta_3 \frac{n}{\tau_2} e^{-\frac{n}{\tau_2}}.$$  \hspace{1cm} (1)

The parameter vector $\phi = (\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2)$ describing the shape of the forward curve is unobserved and will be estimated. This model describes a fixed cross section of forward rates but is straightforward to extend to the dynamic case by allowing $\phi = \phi_t$ to evolve over time. The NSS model extends a prior functional form introduced by Nelson and Siegel (1987) by the additional flexibility of the last term, along with the parameters $\beta_3$ and $\tau_2$.

By using the relation

$$y(n) = \frac{\int_0^n f(n; \phi_t)}{n}$$ \hspace{1cm} (2)

we can transform equation (1) into the following:

$$y(n; \phi) = \beta_0 + \beta_1 \left( 1 - e^{-\frac{n}{\tau_1}} \right) + \beta_2 \left( \frac{n}{\tau_1} - e^{-\frac{n}{\tau_1}} \right) + \beta_3 \left( \frac{n}{\tau_2} - e^{-\frac{n}{\tau_2}} \right).$$ \hspace{1cm} (3)

This describes a curve composed of several additive terms $\beta_0$ through $\beta_3$, with the cross-sectional contribution of each term determined by the shape parameters $\tau_1$ and $\tau_2$. The additive terms $\beta$ can be interpreted in terms of the the Litterman and Scheinkman (1991)
factor decomposition of the term structure into level, slope and curvature, with the shape parameters allowing the nature of slope and curvature to change over time.

The parameter $\beta_0$ is the long-run level of the term structure: all other terms decay to zero as $n \to \infty$. An increase in $\beta_0$ induces a flat increase across the yield curve. Parameter $\beta_1$ is a slope term. Its contribution to the overnight short rate is $y(0) = \beta_0 + \beta_1$. Its loading is monotonically decreasing in $n$, so it is the inverse of a standard slope term. The loadings of the two parameters $\beta_2$ and $\beta_3$ can be described as curvature functions. They they follow a hump-shaped pattern:

$$h(n, \tau) = \frac{1 - \exp(-\frac{n}{\tau_1})}{\frac{n}{\tau_1}} - \exp(-\frac{n}{\tau_1}). \quad (4)$$

The maximum value of this function occurs at $h_{\text{max}}(\tau) = \tau \times K$ where $K \approx 1.793$. With this interpretation in mind, we can describe the NSS model in relation to the Nelson-Siegel model as extending the function form to allow a second flexible curvature component. Many term structures can be fit adequately without the introduction of $\tau_2$, but the second shape parameters allows the model to fit yield curves demonstrating multi-peaked, “S”-like and other forms.

### 2.2 Objective Function

The model is put into action by fitting a cross section of financial instruments. In this paper, we will consider securities with observed prices $P_{i,t}$. We calculate $\hat{P}_{i,t}(\phi_t)$ by discounting the promised cashflows according to the yield curve $y(n; \phi_t)$. Our task is to find the solution:

$$\phi_t = \max_\phi \sum_i w_{i,t} \left( P_{i,t} - \hat{P}_{i,t}(\phi) \right)^2. \quad (5)$$

The choice of weights $w_{i,t}$ embeds certain modeling decisions. In addition to an equal-weighted specification, a common practice is to use the inverse of the Macaulay duration. We will discuss this in the next section. Absence of arbitrage places some restrictions on

---

$^1$We compute the limit as of $y(n; \phi) \to 0$ by applying L’Hôpital’s rule to equation 2, the instantaneous short rate is equal to the forward rate at maturity zero.
plausible model parameters. To enforce the zero lower bound, we require that

\[
\beta_0 > 0 \\
\beta_0 + \beta_1 > 0
\]

so that neither the short rate nor long-dated yields fall below zero. We also require \(\tau_1, \tau_2 > 0\) to rule out an unbounded yield curve. Certain combinations of \(\beta\) and \(\tau\) parameters resulting in negative forward rates can also be ruled out, but in practice this is not a concern.

2.3 Noise Measures

We follow Hu et al. (2013) in estimating yield curve noise measure resulting from the fitted yield curve. At each date \(t\), we calculate the yield-to-maturity \(y_{i,t}\) for the set of \(n\) observed bonds \(B_i\). Then, we calculate the yield-to-maturity that would be derived by using the fitted price \(\hat{P}_{i,t}\) from the estimation above, and denote this by \(\hat{y}_{i,t}\). We define the yield curve noise to be the root mean squared error of these deviations:

\[
\text{noise} = \left( \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_{i,t} - y_{i,t})^2 \right)^{1/2} \tag{6}
\]

The result will depend on the weighting scheme \(w_{i,t}\) used in the objective function. This measure is designed to pair with the inverse-Macaulay duration weighting scheme discussed above - such a scheme is designed to mimic using \(\hat{y}_{i,t} - y_{i,t}\) as the errors to be minimized in the objective function: that is, using the yield-to-maturity of individual securities instead of prices, as advocated by Svensson (1994). The goal of such a specification is to achieve a more accurate estimate of the yield curve at the short end. Because prices of short duration securities are less sensitive to the level of the yield curve, fitting with prices alone may give poor estimates of the yield curve at the short end. For the purposes of monetary policy, yields are the object of interest rather than prices. Weighting bonds by the inverse of their duration achieves a similar effect - prices of short-maturity securities are overweighted and long-dated securities are downweighted. This results in more significant pricing errors at the longer end. However for the purposes of measuring liquidity, conditions, or the possibility

\[\text{recall that the yield-to-maturity for a bond } B \text{ with price } P \text{ is defined to the the unique value } y \text{ such that the cashflows of } B \text{ discounted at the constant rate } y \text{ are equal to the observed price of the bond } P.\]
of arbitrage, it seems more natural to look in pricing space, not yield curve space. To that end I perform an equal-weighting of bonds and then construct a root-mean-square pricing error measure:\[3\]

\[
\text{pricenoise} = \left( \sum_i w_{i,t} \left( P_{i,t} - \hat{P}_{i,t}(\phi) \right) \right)^{1/2}
\]

(7)

3 Estimation

3.1 Data

My primary data source is the CRSP daily Treasury files from January 1991 to December 2017, a sample period including 6749 days. These data primarily derive from the GovPX end-of-day indicative measures, which have some shortcomings as discussed by Jordan and Kuipers (2005) and Adrian et al. (2017). I include both on-the-run and off-the-run issues of non-callable bills, notes and bonds. The prices of short-term Treasury securities display erratic liquidity effects and do not exhibit the same behavior as those of notes and bonds, and the availability of yield curve data is less consistent past the 10 year horizon. For example, Treasury discontinued issuance of new twenty-year bonds beginning on January 1, 1987, and temporarily suspended issuance of 30-year bonds between February 18, 2002 and February 8, 2006. I drop issues maturing outside of a window of 28 days to 10 years from each observation date entirely. This keeps the focus of the analysis on a well-behaved part of the yield curve. I also drop the seasoned notes and bonds which mature in less than 1 year. Finally I drop 6 days with implausible price data from plots and regressions.\[4\] These specifications could be refined even more as many authors have noted specific data errors. I plan to provide more accurate estimates in the future, documented on Github.

The sample contains 2755 unique CUSIPs, including 1557 non-coupon-paying bills, 1089 notes and 109 bonds. On average, there are 155 securities on a given day, though this decreases to a minimum of 78 on May 16, 2003 and increases to around 240 in 2017. I plot two summary statistics in Figure 1.

I use the QuantLib-Python library to estimate daily yield curves. This open source

\[\text{Hu et al.} (2013) \text{ note that an equal-weighted estimates give similar results to noise measure, but here we are studying a different measure which requires an alternative objective function.}
\]

\[\text{These are 1998-01-23, 1998-07-10, 1998-07-21, 1999-01-28, 2006-06-30 and 2013-11-22. The majority are reported in the aftermath of the Asian financial crisis. None of these dates falls at the end of a month; their removal does not affect monthly regression estimates.}\]
Figure 1: Properties of the Treasury data reported by CRSP. The blue line plots the number of securities observed every day and is measured on the right hand axis. The average maturity of the outstanding securities is plotted on the left hand axis. The high frequency variation in the average maturity is primarily due to the monthly auction schedule. The number and variety of securities decreased during a period of United States Government budget surplus.

package provides libraries for bond cashflow analysis, curve fitting and interest rate computations. Treasury securities are specified with actual/actual business day counts, with an end-of-month business day convention, same-day settlement. The Quantlib library contains all necessary data about US Treasury coupon repayment, which I verify by checking that the accrued interest calculated by the library exactly matches the accrued interest reported by CRSP.

To focus the analysis on consistent observations, I construct the noise measures including only Treasury notes and bonds. This focuses the estimation on a window of securities with maturity between than 1 year to 10 years. Leaving shorter maturity bills in the initial estimation is important for model identification and stability.

3.2 Parameter restrictions

Efficient estimation of the model relies on the fact that, conditional on $\tau_1$ and $\tau_2$, bond prices are a linear function of the $\beta$ parameters, so they can be estimated using least squares. This results in a two stage procedure: first, the parameter $\tau$ is chosen. Second, conduct a linear regression. This was recognized by [Nelson and Siegel (1987)], who chose to estimate the model by sampling $\tau_1$ over a fixed interval, without attempting to search for the global maximum. Diebold and Li (2006) estimate a time series of Nelson-Siegel model parameters by fixing $\tau$ in advance at 0.609, placing a hump near the 1-year maturity mark. This allows a simple, direct interpretation of $\beta_i$ estimated on different days as linear factor components driving the yield curve, as fixing $\tau$ means their loadings are time invariant. Unfortunately neither of these approaches are suitable for the purpose of estimating noise. This requires a full optimization: search over the entire sample space for the optimal $\tau_1$ and $\tau_2$, using the optimal $\beta$ for each sample.

The full optimization problem poses several major challenges to estimation in practice.
Gilli et al. (2010) identify two main factors at play. First, the objective function is has significant nonconvexities in general. The second is that the objective function is ill-conditioned for certain values of the shape parameters $\tau$. If $\tau_1 = \tau_2$, the parameters $\beta_2$ and $\beta_3$ are unidentified, as their loadings are given by $h(\tau_1) = h(\tau_2)$. The optimization will be nearly singular if $\tau_1$ and $\tau_2$ are not equal but are numerically close. Another problem is that if $\tau_1$ is very large or very small, with the hump at the origin or very far away from it, then $h(\tau)$ will take the form of a simple downward or upward slope over the observed maturity range of the yield curve. As $\tau_1$ approaches 0 or grows large, the correlation between the factor loadings on $\beta_2$ and $\beta_1$ restricted to a fixed interval $[\epsilon, N]$ approaches ±1. A similar problem occurs for $\tau_2$ and $\beta_3$. This results in large, unstable estimates of $\beta$ parameters over time. Gilli et al. (2010) suggest ranges for $\tau_i$ to prevent this:

$$0 \leq \tau_1 \leq 2.5$$
$$2.5 \leq \tau_2 \leq 5.5$$

implying a $\beta_2$ curvature term peaking between 0 and approximately 4.5 years and $\beta_3$ curvature term peaking between 4.5 and 9.8 years. They also suggest imposing restrictions on $\beta$, requiring that $0 \leq \beta_0 \leq 15$, $-15 \leq \beta_1 \leq 30$ and that the absolute value of $\beta_2$ and $\beta_3$ are less than 30. Imposing such tight restrictions causes significant deterioration in the quality of the estimates. I keep the restrictions $0 \leq \tau_1 \leq 10$ and $\tau_2 > \tau_1$, the latter of which is simple to implement by optimizing over a nonnegative dummy parameter $\xi = \tau_2 - \tau_1$, and find these do not cause a significant impact on fit performance.

### 3.3 Optimization techniques

Addressing the nonconvexities in the optimization problem is a further challenge. Early authors did not attempt to validate whether their optimization results are trustworthy. These days, it is standard to attack this class of problem with heuristic or stochastic optimization methods. Manousopoulo and Michalopoulos (2009) compare BFGS, Frank-Wolfe Reduced Gradient Method, Nelder-Mead, Powell’s Direction Set algorithm, and Simulated Annealing, however their results are inconclusive, generally supporting the latter three methods,

---

*Estimating the model with these restrictions resulted in a 25% increase in mean yield curve fitting error and a 30% increase in the standard deviation of the estimates.*
while gradient-based BFGS and Frank-Wolfe perform poorly. Gilli et al. (2010) recommend Differential Evolution. QuantLib offers a variety of optimization methods which allow me to evaluate these techniques.

Differential Evolution is a heuristic method introduced by Storn and Price (1997). It is characterized as an evolutionary-class algorithm in which candidate points are moved according to vector differences from the current best candidate. Although this method has a few simple tuning parameters, I did not find it an effective solution method. Using Gilli et al. (2010)'s recommended parameter choices resulted in slow convergence and long wait times. Adjusting tuning parameters did not make this method any more attractive. I would like to review the source code of the implementation before I state this definitively, but a quick glance makes it seem unlikely to be the problem. I believe the main difference is that their estimation uses the Diebold and Li (2006) dataset of precomputed zero rates; this computationally demanding method seems less suitable applied to the far heavier CRSP dataset.

Like Manousopoulos and Michalopoulos (2009), I find the well-known Simulated Annealing algorithm to perform reasonably well. However, this was beat by a simpler technique: a repeated Nelder-Mead search with randomized starting points. When using a local optimization method, the solution depends on the starting value - this applies to Simulated Annealing as well as Nelder-Mead. Executing a Nelder-Mead search to convergence takes under 1 second on an Intel i7-8700K, so randomizing over the parameter space is very straightforward. I begin my search using the estimates of Gurkaynak et al. (2007) published by the Federal Reserve Board, and then add random normal jumps with unit standard deviation to my parameters, keeping them in bounds, and perform five such searches. Once I have passed through the dataset, I loop through again using the previous estimate plus noise, as well as the estimates from the previous two days. This method is simple to implement and can be parallelized easily. I iterate through the dataset until no improvements are found at any day in the sample; with over 6000 observations I am confident this results in estimates close to the global optimum. Splitting this estimation window into 10 subperiods estimated in parallel results in convergence overnight.

Annaert et al. (2013) propose another method of estimating the Nelson-Siegel model to obtain more consistent time series estimates, overcoming the collinearity described above:
they introduce $L_2$ shrinkage term, which takes the form of a ridge regression in the second stage. To capture this, I construct weighted estimates of the Nelson-Siegel yield curves applying a small penalty of $\lambda = .001$ times the sum of squared parameter values, inclusive of $\tau$, in the objective function. They demonstrate this technique produces a marked increase in the time series stability of the parameter estimates. Like other regularization techniques, this also results in better out-of-sample performance in the cross section. I implement this as a second benchmark for two reasons: first, extreme parameter estimates and unusual shapes in the yield curve may be indications of noise that a naive estimate fails to capture. Second, the improved numerical stability may result in more reliable optimization results.

### 3.4 A comment on the visual accuracy of par yield curves

The yield curve is a visual tool useful for summarizing price data of many differing securities in a one-dimensional form. The purest theoretical form does not reflect the properties of securities that are accurately traded. The par yield curve is an alternative visualization meant to do a better job.

![Figure 2: Fitted par yield curve versus fitted prices on May 19, 1995.](image)

The semiannual par yield at maturity $n$ observed at time $t$ is defined as the yield-to-maturity of a theoretical bond with semiannual coupon payments maturing at time $t + n$. If $n$ is not a precise multiple of 6 months, the first coupon will be paid at some time $t_0$ sooner.

---

6Their method does not use a fixed penalty, but switches between constrained and unconstrained estimates, only imposing the penalty term when the second stage $\beta$ regression is collinear, as indicated by the condition number.
than 6 months, and the bond will carry accrued interest. If \( d_t(n) \) is the discount function at time \( t \) for a risk free payment at time \( t + n \):

\[
y^p_t(n) = \frac{2(1 - d_t(n))}{\sum_{i=1}^{2^n-t_0=n} d_t(i/2 + t_0) - a_t(n)}
\]

(8)

where \( a_t(n) \) is the fraction of accrued interest on the hypothetical bond, a function approaching 1 at the time of each coupon payment. Annual and quarterly par yield curves can be defined for bonds with other payment frequencies. This makes for an adequate comparison to the yield-to-maturity of a traded bond, as both correspond to weighted averages of zero-coupon interest rates.

An ideal par yield curve will align nicely with the bonds right along the line, which is not what we see in Figure 2. Here there is significant dispersion, especially at the long end. A naive interpretation is that this is the object we are measuring? In fact not: the average yield curve noise on this day only 2.01 basis points. This is displayed on the right hand graph. The reason older securities fall below the curve is that they were issued during a time of falling interest rates: their larger coupons relative the prevailing yields decrease the effective duration of these securities, and the one-dimensional maturity axis cannot capture this. The par yield curve is a useful object when graphical intuition is necessary or when curves are imprecisely measured, but they are less appropriate for identifying issues in high-precision fitting; pricing errors should be used instead.

4 Results

4.1 Estimated yield curves

Another basic assessment comparing the performance of the optimization methods is what sort of yield curves they estimate. In what follows I will be referring to my four techniques under comparison: the standard duration-weighted measure used as a benchmark, the ridge-regression approach Annaert et al. (2013), and then equal-weighted versions of both. In figure 3 I plot estimated yields from the four techniques under consideration, along with the estimates provided by the Federal Reserve Board described in Gurkaynak
et al. (2007). Figure 3 visually shows that, compared to the time series variation, the estimates in methodologies produce small changes in the yield curve. This is a reassuring picture to those who are attempting to fit yield curves - all methodologies generally agreed within a few basis points.

Figure 3: Time series of zero coupon yield estimates. GSW is the yield curve described by Gurkaynak et al. (2007) published by the Federal Reserve Board.

Recall that to estimate these curves, we have focused on bonds with maturities between days month to 10 years, and have included the on-the-run and off-the-run issues. This leads to a discrepancy between the results of the four models under consideration and the GSW yield curve, which is fitted to a broader maturity spectrum of Treasury securities, including maturities out to 30 years. In addition, GSW drop on-the-run and first-off-the-run issues from their estimation, focusing on an estimating an off-the-run yield curve. These selection choices are reflected more significant difference in the estimates at the ten-year horizon: our estimates are uniformly lower, reflecting the model’s attempt to fit the low-yield on-the-run issues.

Figure 4 plots a measure of the degree to which the four estimates differ, making the differences more clear than in a time series graph. At each date I compute the average yield curve \( \bar{y}_t(n) = \frac{1}{4} \sum y^k_t(n) \) from the four estimation methods indexed by \( k \). I then compute the root mean square deviation of \( y^k_t(n) - \bar{y}_t(n) \) for each maturity over the time series. Overall, the agreement is close, and a deviation of .5 to 1 basis point for different measures may be acceptable depending on the application. However, the yield curve noise estimate is also on the order of basis points, so for our purposes this does not demonstrate they are
equivalent. One interesting behavior is the increasing deviation near the tail end of the yield curve. I interpret this as overfitting due to the lack of long-maturity securities included in the estimation. If time permitted, I would conduct an estimation including securities out to 15 years to form more accurate estimates over the current window.

4.2 Noise Measures

Figure 5 contains time series plots of four variations of Treasury liquidity measure. The first panel compares two duration-weighted estimates of yield curve yoise, computed by an unconstrained optimization and a regularized version. The second panel compares two equal-weighted measures of mean squared pricing error.

The average value of the noise measure over the sample period is 3.12 basis points using the unconstrained estimator, with a standard deviation of 1.96 basis points. For the equal weighted pricing error, the sample average is $0.122 per $100 face value, with a standard deviation of $0.126. One remarkable development is how much the yield curve noise has decreased in recent years. In the sample period from January 2011 to December 2017, the average noise decreased to 1.73 basis points with a standard deviation of .40, and the average price error decreased to $.064 per $100, with a standard deviation of only $.017. The Treasury market is functioning near historic highs of liquidity according to this measure.
As seen in the second panel of Figure 5, the results of the ridge regression technique are remarkably similar to the standard estimates. However this method drastically reduces the variance of parameter estimates. In contrast to imposing hard parameter constraints, the average fit with a regularized model is very good. The average fit error increases by only .04 basis points on average, or 1.3% of the baseline estimate, and as recorded in Table 1, the correlation between the two measures is over 99%. However, the standard deviation of the estimates of each of the $\beta$ parameters decreases by over 99%. Using regularization on the equal-weighted price error measure is similarly effective.

Taking into account the increase in the number of securities as seen in Figure 1, the Treasury market in recent years appears to be functioning smoothly with no mispricing comparable to the 2000s dot-com bust or the 2008 financial crisis. However, this extended sample shows that the noise measure continues to be an informative marker of financial market developments. Figure 6 plots estimates of the yield curve beginning in January 2011, with several upticks labeled with concurrent episodes of financial market turmoil. Of particular note is the sensitivity of the Treasury market to political risk - the yield curve noise demonstrates peaks during both debt ceiling crises in 2011 and 2013, and spiked up again the lead-up to the government shutdown in January 2018.

In Figure 7 I compare the benchmark estimate to two other measures. Motivated by the pricing relation:

$$P_t^{(n)} = \exp(ny_t^{(n)})$$
we might expect that the so one final potential estimator is taking the log of the price noise. To make this comparison, I plot the log price noise and yield curve noise demeaned and standardized in the left hand panel of Figure 7. The new measure appears visually similar to the original noise measure, with the same time series properties. The process increases the variation in the price measure during normal times and reduces the spike during the financial crisis.

Finally, I compare my estimates to the yield curve noise estimated by [Hu et al. (2013)].
Overall the estimates are quite similar, with a daily correlation of 95%, as recorded in Table 1. Neither estimate is uniformly larger or smaller, though my estimate of the yield curve noise is slightly higher during the 1997 Asian financial crisis and the dot-com bust and slightly lower during the 2008 financial crisis. Overall it seems the differences due to data cleaning and processing are as significant as the algorithm used.

<table>
<thead>
<tr>
<th></th>
<th>Noise</th>
<th>Noise reg.</th>
<th>HPW</th>
<th>Price err</th>
<th>Price err reg.</th>
<th>Log price err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>1</td>
<td>0.996</td>
<td>0.954</td>
<td>0.878</td>
<td>0.876</td>
<td>0.837</td>
</tr>
<tr>
<td>Regularized</td>
<td>1</td>
<td>0.953</td>
<td>0.876</td>
<td>0.874</td>
<td>0.837</td>
<td></td>
</tr>
<tr>
<td>HPW</td>
<td>1</td>
<td>0.949</td>
<td>0.948</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price err</td>
<td></td>
<td>1</td>
<td>0.999</td>
<td>0.855</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price err reg.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>0.851</td>
</tr>
<tr>
<td>Log price err</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 1:** Correlation between yield curve noise estimates.

### 4.3 Liquidity measures

Another evaluation of the liquidity measures is their informational content compared to other common measures of liquidity. Adrian et al. (2017) include a version of the measure computed by the Federal Reserve Board as one of their benchmark liquidity factors. Here I estimate the extent to which the liquidity factor reflects this information as well. One use of the yield curve fitting error is as one of a benchmark of many liquidity measures. Bisias, Flood, Lo & Valavanis (2012) suggest that quantifying a complex notion like the health of the financial system requires taking advantage of many indicators from varied sources and helping to contribute to this goal is a motivator for providing an open-source implementation of this construction.

Table 2 displays the correlations between my estimates and other variables associated with Treasury pricing and broader financial market conditions. I take first-differences in all estimations to remove time series trends. All variables match the 324 monthly observations over the 27 year sample period except GCRepo, which starts in 1998. The first set of measures are standard constant-maturity Treasury yields from the Federal Reserve Board’s H.15 release and obtained through FRED. $y_{3m}$ is the three-month constant maturity rate. Slope is the ten year slope of the yield curve, which is the ten-year yield minus the one-year
yield. The third measure is a 21-day rolling average of the five-year constant maturity yield. Overall the most important measure among these conditions is the yield curve slope. This is consistent with the slope as a measure of risk premium in the Treasury market. The three month yield and bond yield volatility are not significant in univariate regressions and the signs are indeterminate. The bond volatility measure does not vary by much over time, remaining close to its sample average of .36 percentage points annualized. Treasury prices remain relatively stable even in periods of stress. This suggests the yield curve fitting error is not related to simple changes in market prices.

The next two measures are $On_{5y}$ and $On_{10y}$, the on-the-run premium for the five-year and ten-year Treasuries respectively. These are the yield differences of these securities relative to the yield curve, which is included as a component of the yield curve noise, and I measure it using the same fitted values. The high levels of correlation here should be taken with a grain of salt. The on-the-run premium for the 10 year bond matches very closely with the overall level of the spline fitting error, as the ten-year yield is relatively poorly pinned down in our estimates, and times when overall fit is poor this estimate of the , and in fact the correlation in levels instead of first-differences is approaching 80%. Regardless, the estimates of these measures line up well with other estimates of the on-the-run premium, and suggests that the premium to holding liquid securities is especially valuable when there is disagreement in pricing.

Repos the General Collateral Repo rate published by the Federal Reserve Bank of New York. This measure is typically closely tied to the Federal Funds rate and captures similar information to the three-month Treasury yield. Although the repo rate is impacted by liquidity events, the $R^2$ in a univariate regression is only on the order of 2%. Libor, obtained from FRED, is very similar measure closely tied to the short rate.
The final variables measuring broader market conditions are $Baa - Aaa$, the default spread of $Baa$ over $Aaa$ rated seasoned corporate bonds obtained from FRED, which also makes available $VIX$, the value-weighted market return published by Kenneth French, and the Pastor and Stambaugh (2003) measure of market liquidity. Of these, the first three have significant explanatory power. In particular, the default spread, which spiked up during the financial crisis, seems more closely related to the noise measure than any of the other variables.

The $R^2$ of a combined regression of the yield curve noise on all the other measures is 46%, and excluding the 10-year on-the-run premium this drops to 33%. The related figures for the price error are 62% and 39% respectively. This demonstrates that yield curve noise contains unique information about financial market conditions. Overall the price error has higher correlations and more of its variation can be explained, but this is due to the fact that the crisis period is a more significant component of this measure’s variation.

4.4 Pricing Currency Portfolios

Another use of the Treasury liquidity measure is in asset pricing. Lustig et al. (2011) document that currency returns are driven by a global factor related to equity market volatility, and Du et al. (2017) document the persistent effects of liquidity conditions and the lack of arbitrage capital on currency prices. We would like to understand the degree to which the noise measure can help explain this factor.

To this end I estimate pricing regressions of the form

\[ R_{i,t} = \alpha_i + \beta_{noise_k} \Delta Noise_{k,t} + \beta_{Mkt} Mkt_t + \epsilon_{it} \]

where $Mkt_t$ is the US equity market return and $Noise_{k,t}$ is indexed among $K = 5$ yield curve noise estimates, which includes the four measures above and the log of the equal-weighted price measure. I display the results in Table 3. Portfolio 1 and Portfolio 6 the broad-market returns on portfolios of currencies sorted by interest rate, made available by Adrien Verdelhan. Portfolio 1 is comprised of low interest rate countries and Portfolio 6 contains high interest-rate countries. Overall, the noise measure has significant explanatory power for movements in high interest rate countries, which are more susceptible to financial
turmoil. $\beta_{\text{noise}}$ is highly statistically significant for all measures. They are less successful explaining movements of low-interest rate countries. Overall the results for the different methodologies are very similar, however the log Price seems to lose most of the explanatory power found in the other measures.

<table>
<thead>
<tr>
<th></th>
<th>Portfolio 1</th>
<th></th>
<th>Portfolio 6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_{\text{noise}}$</td>
<td>$\beta_{\text{MKT}}$</td>
<td>$R^2$</td>
<td>alpha</td>
</tr>
<tr>
<td><strong>Benchmark</strong></td>
<td>-0.369</td>
<td>-0.023</td>
<td>2.38%</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>[-2.62]</td>
<td>[-0.82]</td>
<td>[-3.11]</td>
<td></td>
</tr>
<tr>
<td><strong>Regularized</strong></td>
<td>-0.401</td>
<td>-0.023</td>
<td>2.82%</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>[-2.87]</td>
<td>[-0.83]</td>
<td>[-3.12]</td>
<td></td>
</tr>
<tr>
<td><strong>Price (EW)</strong></td>
<td>-6.142</td>
<td>-0.027</td>
<td>2.09%</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>[-2.44]</td>
<td>[-0.96]</td>
<td>[-3.06]</td>
<td></td>
</tr>
<tr>
<td><strong>Reg. Price (EW)</strong></td>
<td>-5.959</td>
<td>-0.026</td>
<td>2.04%</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>[-2.41]</td>
<td>[-0.92]</td>
<td>[-3.06]</td>
<td></td>
</tr>
<tr>
<td><strong>log Price</strong></td>
<td>-0.077</td>
<td>-0.014</td>
<td>0.99%</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>[-0.14]</td>
<td>[-0.49]</td>
<td>[-3.10]</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3:** Currency portfolio pricing results. The benchmark model includes the end-of-month estimates of yield curve noise and the Mkt return.

### 5 Conclusion

This paper should serve as a useful guide for anyone attempting to fit this class of models in the future. There are too many open ends to reach a definitive conclusion about the best practice for estimating this model. As I prepare my contributions to QuantLib, I would like to give the Differential Evolution algorithm a closer inspection to see if I can determine the reasons it was so problematic. There are other measures of liquidity I would have included in my analysis save for time constraints, including Longstaff (2004)’s Refcorp spread, the Libor/OIS spread, and the microstructure-based liquidity measure of Adrian et al. (2017).

Another straightforward extension is to include longer-dated securities in the estimation. This does not come without problems, as it involves a new class of modeling choices and exclusions. However, we have seen that the long end of the yield curve is not well identified in this procedure, and it may be worth the effort. Another point this raises is the gap in the
term structure: Treasury’s hold from issuing 30-year securities in the early 2000s has left no long term bonds maturing in the years 2032-2036. Understanding the fit and behavior of the yield curve with respect to this gap could prove fruitful.

Another direction is to extend the class of models under consideration. One that would be quite straightforward would be the ridge regression with time-varying penalty proposed by Gilli et al. (2010). This would compare naturally against the simpler version used in this paper. Another possible modeling choice is to choose a weighting scheme $w_{it}$ that redistributes weight across various parts of the yield curve according to the density of observations. On a given day there are typically many more securities with maturity less than five years than there are longer-term bonds. This also interacts with the existing weighting scheme, giving more importance to the shorter end of the curve. Computing this and producing a similarly reweighted noise measure would generate an estimate of the average error across the yield curve. Finally, Diebold and Rudebusch (2013) demonstrate that the NSS parameterization is inconsistent with no-arbitrage as it does not account for a price of risk. They demonstrate afterward that a NSS model extended with a volatility matrix is compatible: their “arbitrage-free Nelson-Svensson-Siegel” model could be a good candidate for further refining yield curve estimation. This model introduces additional parameters, can like the NSS model it be simplified into a search over the shape and volatility parameters, $\tau$ and $\Sigma$, followed by a linear regression.

Extending this framework beyond CRSP is a natural next step. Jordan and Kuipers (2005) find that GovPX and FRBNY data differ by a significant amount, especially in the pricing of notes and bonds. Unfortunately GovPX intraday data do not contain a broad enough cross-section to construct an accurate comparison measure. Abrahams et al. (2016) find the error estimates computed by the Federal Reserve Board are a more accurate liquidity proxy than estimates computed from CRSP in an affine term structure model. The curve-fitting framework I provide could also be used to probe the information contained in intraday yield curve evolution.
Bibliography


Polychronis Manousopoulos and Michalis Michalopoulos. 2009. Comparison of non-linear


