Stockholm School of Economics

Department of Finance

Master Thesis

Spring 2014

Empirical Risk Decomposition of Break-even Inflation

Jonas Gustavsson†   Patrik Paluch‡

Abstract: Using principal component analysis, we find that there are three principal components that can explain almost all of the variation in the 5-20 year maturities of break-even inflation, defined as the spread between nominal Treasury and TIPS yields. The three principal components correspond to the factors regularly used in the empirical literature to describe the term structure of interest rates, namely, level, slope and curvature. The first principal component, explaining over 90 percent of the variation in the data, is very similar to 10-year break-even inflation. We regress both of these series on variables aimed at embodying different sources of risk relevant to break-even inflation, and find that short-term inflation expectations and differences in liquidity have the highest explanatory power. In accordance with our theoretical discussion of risk premia, we also include measures for the inflation risk premium and volatility risk. Although these measures do not add much to the analysis of the first principal component, they contribute significantly to the explanation of the second principal component. We interpret this as an indication that the inflation risk premium is an important driver of the slope of break-even inflation. We also find that changes in liquidity conditions impacts the slope of the term structure, but propose that this merely reflects a change in short-term inflation expectations.

Keywords: TIPS, Break-even inflation, Term structure, Liquidity, Inflation

Tutor: Irina Zviadadze, Stockholm School of Economics

† 22038@student.hhs.se

‡ 21918@student.hhs.se

TABLE OF CONTENTS

1 Introduction

Following the example of the United Kingdom, Canada, Sweden and New Zealand, the United States introduced their version of inflation-indexed debt in 1997 called Treasury inflation-protected securities (TIPS). During its infancy on the US debt market, the investor participation was low, as was trading volume. Perhaps this was due to a lack of familiarity, or because of rules preventing institutions from holding such instruments. In May 2001, the Treasury Borrowing Advisory Committee of the Bond Market Association recommended the Treasury to stop issuing debt linked to inflation.[1] Given their later success, the welfare benefits from the decision of continued issuance have probably been higher than they would have been without it. The popularity of the inflation-protected bonds has since then increased greatly, from an estimated 2% of total outstanding US government debt in January 1999, to 8% in January 2014.[2]

Previous studies like Evans (1998), Pflueger and Viceira (2011), and Ilmanen (1995) have tested and found predictability in both inflation-indexed and nominal bond excess returns. The two main competing explanations for observed return predictability are time-varying risk premia and market inefficiency. Time-varying risk premia may reflect either assets' varying riskiness or investors' changing risk aversion. In the empirical literature, these two are often not separated, and in light of this, we will focus much of this thesis on examining various risk premia in the context of break-even inflation. We define break-even inflation as the spread between nominal US Treasury bond yields and TIPS yields in the cash market.

It has been established in the literature that break-even inflation can be mainly decomposed into expected inflation and a liquidity premium. To name a few important contributions to the research on this topic, D'Amico, Kim and Wei (2010) estimate the liquidity premium using three different measures of liquidity. By employing surveys of forecasted inflation, they also take expected inflation into account. While Grishchenko and Huang (2012) use mostly the same liquidity variables, they instead focus on estimating the inflation risk premium. Pflueger and Viceira (2011) add another aspect of risk to the discussion when devoting some attention to the real interest rate risk premium. In Pflueger and Viceira (2013), they build upon their earlier study by estimating time-varying risk premia for the real rate, inflation and liquidity.

We examine to what extent break-even inflation can be explained as compensation for various sources of systematic risk by first identifying relevant bond risks, and then assigning suitable proxies for them. By using these proxies to explain the variation in break-even inflation,

[1] From a report to the Secretary of the Treasury. See references for a source available online.
[2] According to TreasuryDirect Monthly Statements of the Public Debt, total (marketable) debt outstanding.

we effectively find detailed information about the market prices of risk and the different risk factors' relative importance to break-even inflation.

In contrast to other studies on the subject, which mainly focus on the difference between 10-year nominal and inflation-indexed bonds, we will focus on the term structure of break-even inflation. To the best of our knowledge, no previous study has aimed to discern how well a broad collection of risk components explain break-even inflation across different maturities. Whereas other authors using the term structure mainly focus on a specific risk premium (such as the inflation risk premium), we attempt to determine how much of the variation in the data that can be explained by a wider array of risk proxies while controlling for both short- and long-term expected inflation.

For inflation expectations over the longer-term we utilize inflation forecasts published by the University of Michigan while the Chicago Fed National Activity Index is used as an indication of expected inflation over the next year. The other variables which we make use of in our analysis represent liquidity, volatility and inflation risk. To capture different characteristics of liquidity risk, we begin by employing the off-the-run spread. This spread is the difference in yield between off-the-run and on-the-run nominal Treasury bonds commonly used in gauging the "flight to liquidity". We also use the relative Z-spread between asset-swapped nominal Treasury bonds and asset-swapped TIPS which takes current and anticipated future differences in costs of financing into account. Furthermore, the ratio of TIPS to nominal Treasury transaction volumes is utilized, and lastly, the turnover ratio of TIPS transaction volume to the amount of TIPS outstanding. For the volatility and inflation risks, we use the Merrill Lynch Option Volatility Estimate Index and the volatility of the CPI-U, respectively. A considerable part of this thesis will be dedicated to carefully describe and evaluate why, and to what extent, these measures are appropriate for explaining break-even inflation.

By decomposing the yield data into different risk measures, we will more plainly be able to see if different sources of risk are discounted by the market, and in what way they impact break-even inflation. We find that the term structure of break-even inflation rates can be summarized by three principal components corresponding to the level, slope and curvature factors that are so well-known in the term structure literature.

Our results show that the first principal component, which is equally weighted on all the maturities, explains as much as 93% of the variation in 5-20 year break-even inflation rates. Controlling for short- and long-term expected inflation, we find that our proxies for liquidity are both statistically and economically significant. This suggests that caution should be exercised by those who want to approximate market-based inflation expectations with the help of break-even inflation. For the full model, with variables aimed at embodying volatility and inflation risk included, the R^2 reaches 64%. In terms of economic significance, the relative Z-spread proves to be an important gauge as a liquidity element for the first principal component. The variable for short-term inflation expectations is the second most important variable for explaining level changes in the term structure.

For the 10-year break-even inflation, the most noteworthy proxies, not only in terms of statistical significance but also economically, are the short-term expected inflation, the off-the-run spread and the relative Z-spread. An increase of 20 bps (roughly one standard deviation) implies a 13 bps drop in the 10-year break-even inflation, all else equal.

Although our measure for volatility risk is shown to have a high degree of co-movement with the off-the-run spread (which demands caution when making qualitative interpretations of the results), it is still able to perform quite well. This supports the case for including volatility risk into analyses of break-even inflation.

Even though our inflation risk proxy contributes with a large amount of explanatory power only to our second principal component, it is still statistically significant at a 5% level (or less) for all three principal components. We interpret this result as evidence for the existence of an upward sloping term structure of inflation risk premia in nominal Treasuries, consistent with earlier research on this subject. For the second principal component, neither short- nor long-term expected inflation seem to matter much, though the liquidity components once again might contain important information. This is most likely due to the large effects that tightening liquidity has on short-term TIPS yields via the CPI.

While no evidence of multicollinearity between our regressors is found, we have followed this up by testing for serial correlation in regression residuals (a possible weakness we believe is often overlooked) in order to thoroughly assess the validity of our results. We find strong evidence that such autocorrelation in the regression residuals exist, however, it is not persistent (it dissipates within a few lags).

This study could be interesting to both academics and practitioners (such as asset managers) as a way to underscore the intricacy of market data and as an indication of the relative importance of different components in explaining break-even inflation. It also hints at the importance of models that take the various underlying components into account in order to obtain more accurate market-based measures of for example inflation expectations. Our approach could therefore be of interest to, for example, bond fund managers seeking to separate exposures to different systematic risk factors, or to economists attempting to measure inflation expectations by looking at market data.

The thesis is arranged as follows. Section 2 provides background information on the inflation-indexed bond market and bond market risks. Section 3 presents the relevant previous research on the subject. In Section 4, we discuss and describe the data used in the study. Section 5 describes the methodology. Section 6 presents our results and analysis. Section 7 concludes, and Section 8 comprises a few ideas for further research.

2 Background on inflation-indexed bonds

2.1 Indexed securities

An indexed security is an instrument whose cash flows are linked to an index of some kind, such as a basket of interest rates or commodities. One particular class of indexed securities is inflation-indexed bonds. These bonds can help provide investors with means of protecting their real wealth by being linked to changes in the general level of prices in the real economy. An investor in an inflation-indexed bond will with certainty know the real return to maturity on the bond.[3] However, if cash flows are taxed on a nominal basis, inflation risk is reintroduced, and real yields after taxes become uncertain. The nominal return on indexed bonds will only be known ex post though, because the cash return will depend on the future movements in the price index. Thus, the situation is reversed compared to nominal bond returns. In other words, the nominal rate of return to maturity on a nominal bond is certain, while the real return is uncertain, because future fluctuations in the rate of inflation are unknown ex ante.

As written in Deacon, Derry and Mirfendereski (2004), most indexed bonds are linked to a domestic Consumer Price Index (CPI), but measures like the GDP deflator, wholesale prices, and gold prices have also been used historically. The most common type of inflation-indexed bond by far is the capital-indexed bond (CIB). CIBs and indexed bonds in general can be described as consisting of two parts: a real rate of return and compensation for the erosion of the purchasing power of money due to inflation. CIBs have a fixed real coupon rate and a nominal principal which is adjusted for inflation, and used to compute cash coupon payments and accrued interest. Nominal coupon payments on the most widely traded CIB, TIPS, are given by

$$\frac{c}{2} \times \frac{I_t}{I_0} \times M$$

and the principal payment is given by

$$M \times \max\left\{\frac{I_T}{I_0}, 1\right\}$$

where c is the annual real coupon rate, I_t is the value of the price index at the coupon date, I_0 is the value of the price index at the issue date of the bond, I_T is the value of the price index at the

[3] Assuming there are no risk premia, convexity effect, reinvestment risk, and that the price index accurately captures all the relevant price movements in the economy.

maturity of the bond, and M is the face value of the bond. The principal is adjusted semi-annually and includes a "deflation floor"[4] which ensures that the principal payment will always be the higher of the initial face value and the inflation-adjusted principal.

In the case of TIPS, the price index used is the non-seasonally adjusted Consumer Price Index for All Urban Consumers (CPI-U) published by the Bureau of Labor Statistics. The CPI-U measures the average monthly change in the prices of goods and services paid by consumers in urban areas, which includes about 87% of the total US population. The CPI-U is based on prices for food and beverages, housing, clothing, transportation, medical care, recreation, education and communication, and other goods and services. Prices of about 80,000 items are collected each month, and 14,000 families across 87 urban areas provide information on where the households purchase their goods and services. The index is calculated by averaging the price changes with weights proportional to their importance to the spending of the group in question.[5]

The daily reference index, CPI_t, is used as the price index for TIPS. It is a linear interpolation between the second and third month's lagged CPI, and is computed according to the formula

$$CPI_t = CPI_{m-3} + \frac{d-1}{D_m}(CPI_{m-2} - CPI_{m-3})$$

where CPI_{m-3} is the three month lagged CPI-U statistic, d is the day of the current month, subscript m is the current month, and D_m is the number of days in the current month.

As hinted at above, if we simplify, real bonds can be said to consist of a real rate of return and compensation for inflation. The real rate of return to maturity on an indexed bond at the time of purchase can be seen as being certain under some specific assumptions. In practice though, indexed bonds cannot offer complete certainty with regard to a real rate of return. One reason for this is that the price indices used to adjust cash flows will most likely not entirely correspond to a given investor's consumption basket. Only if the investor consumes exactly the right proportions of the exact same goods and services that are included in the price index can we say that the investor can be certain that the cash flows from the indexed bond protects him from declines in purchasing power.

A second reason for why there cannot be complete confidence about real yields is that there is a time lag in the price indices of these bonds. It is simply not possible to publish

[4] The deflation floor can be seen as an option inherent in TIPS, but for simplicity we will assume that it takes the value 0.

[5] For more information about the CPI-U, please consult the source in the references.

continuously updated price indices. Therefore, there will always be a delay between the time for which a price index is used, and the time at which it is calculated and released. For example, you always have to go through the process of collecting price data before publishing an index. Inevitably, this means that inflation numbers that are not entirely up-to-date must be used, and as a result, investors will not be protected against increases in inflation occurring at the end of a bond's life. Thus, the longer an indexation lag is, the poorer the inflation protection will be.

Indexation lags vary quite a lot between different markets. For instance, UK inflation-linked gilts previously used the Retail Price Index (RPI) with an 8-month lag, but changed to using a 3-month lag in 2005. TIPS are linked to the CPI-U with a 3-month lag, as are most other international inflation-linked bond issues, following the method in Canadian RRBs.

A third reason why indexed bonds do not deliver entirely certain real yields is because of potential tax issues. Indexed bonds will generally have a higher duration than nominal bonds with a similar cash flow structure, although nominal rates are more volatile than TIPS rates in normal market conditions. The reason for the higher durations of indexed bonds is that the nominal cash flows to an indexed bond will grow over time, generating higher and higher coupons and principal as time goes by (assuming a positive rate of inflation). This makes inflation-indexed bonds a good hedge for investors with liabilities with high duration. However, it is not entirely correct to directly compare durations between real and nominal bonds. Real bond duration is a measure of price sensitivity with respect to real yields, while nominal bond duration measures price sensitivity with respect to a combination of real yields and inflation expectations.

Another important aspect of price indices worth mentioning is that they often exhibit seasonality. For example, in many industrialized countries consumer spending peaks in the run-up to Christmas, and this is often followed by seasonal price discounts in January. This seasonality will of course impact the prices of indexed debt, especially debt with short maturity. It is therefore important to keep these seasonal swings in mind when pricing such bonds, as the seasonality will affect both nominal coupons and ultimately implied real yields.

The interested reader should consult Appendix B for further historical background on indexed securities, including an exposition of historical real asset returns. In addition, there are sections in Appendix B that touches upon the subjects of Treasury auctions and trading, as well as TIPS pricing.

2.2 Inflation derivatives

Inflation derivatives are to a large extent used to more precisely meet a particular investor's or issuer's demands. Derivative contracts can be used to hedge a multitude of risks, to match the timing and frequency of other cash flows, for index and maturity matching, and so forth. They are also favored by speculators due to the cheap leverage that these instruments provide. The total notional amount of outstanding OTC interest rate swaps as of December 2013 was $461 trillion according to data from the Bank for International Settlements.[6]

One of the most common and basic inflation derivative is the zero-coupon inflation swap. Cash flows are only exchanged at maturity, and involve the exchange of the notional multiplied by the fixed rate (reflecting future expected inflation), and the notional multiplied by the realized inflation. One derivative we will use in this thesis is the asset swap, and a brief introduction to this instrument based on O'Kane (2000) follows.

An asset swap is a derivate instrument which allows the buyer to swap the fixed rate payments on a bond for a floating rate plus an asset swap spread (ASW), this while maintaining exposure to the bond's credit risk. The swap market was born in the early 1990s. It is widely used by banks for hedging floating rate liabilities, e.g. for swapping fixed rate assets for a floating rate to match liabilities like short-term interest on bank deposits. The most widely traded variation of asset swap is the par asset swap. The buyer of the par asset swap buys a fixed rate bond from the asset swap seller for par value. The buyer pays a fixed rate to the seller corresponding to the fixed coupon rate on the bond, and in return receives payments of LIBOR plus the ASW. The ASW is set so that the value of the swap is zero at inception. For this to be achieved, the value of the sale of the bond (par), and the value of the swap transaction has to be equal. Note that interest rate risk is eliminated as the fixed payments from the bond are swapped for floating payments. The buyer still retains the credit exposure to the bond, but will receive appropriate compensation for this. The maturity of the swap is the same as the maturity of the underlying bond.

2.3 Bond risks

2.3.1 Introduction

As with all other financial instruments, there are several sources of risks associated with investing in bonds. Based on Fabozzi (2007), we will describe and comment on some of these in the subsections below, but first, we will emphasize the risks most relevant to spreads between nominal and real bond yields (break-even inflation). These risk factors will form a framework which will help us to structure our analysis of break-even inflation better.

Given exposure to break-even inflation, we postulate that the risk factors that the market requires compensation for can be viewed as being the following:

[6] See references for available source online.

$$BEI = y_{\$} - y_{TIPS} \approx l_n - l_r + \pi + \rho + v_n - v_r$$

where *BEI* is the break-even inflation rate, $y_{\$}$ is a nominal US Treasury bond yield, y_{TIPS} is a TIPS yield (on a TIPS with the same maturity as the nominal bond), $l_n - l_r$ is the difference in compensation for liquidity risk between nominal bonds and TIPS, π is the compensation for expected inflation, ρ is the compensation for inflation risk, and $v_n - v_r$ denotes the difference in compensation for volatility risk between nominal bonds and TIPS. Please see Figure 9 in Appendix A for a look at the term structure of TIPS yields and break-even inflation. Also, note that the real rate and real interest rate risk premium do not affect break-even inflation since both should have the same value in either nominal or real yields[7]. For a more extensive discussion on interest rate risk and other bond risks, we refer readers to Appendix B.

2.3.2 Inflation

A very important component of the break-even inflation rate is expected inflation. In theory, market participants will require some compensation for the expected rate of future inflation when buying nominal bonds as a way to preserve purchasing power. A version of the famous Fisher equation states that a nominal interest rate i can be decomposed into a real rate r, expected inflation π^e, and an inflation risk premium ρ as follows:

$$i = (1 + r)(1 + \pi^e)(1 + \rho) - 1$$

The inflation risk premium is the compensation market participants require for taking on the risk that future unexpected changes in the rate of inflation, above the expected inflation rate, will be realized. This risk premium can in theory be both positive and negative. A positive inflation risk premium implies that the market believes that there is a substantial risk that future realized inflation will turn out to be higher than what is currently expected. In other words, the inflation risk premium can be seen as a measure of the uncertainty about the current inflation numbers that are discounted in the market. A negative inflation risk premium implies the opposite, i.e. that the market believes that there is probably more downside than upside risk to its present inflation expectations.

Campbell and Viceira (2001) estimate an inflation risk premium of 35 bps for 3-month T-bills, and an inflation risk premium of 1.1% over a 10-year horizon for data over the period 1952-1996. Chernov and Mueller (2011) suggest that inflation risk premia can be both positive and

[7] Here, we have to assume that the market accepts the CPI-U as a measure of the "true" rate of inflation.

negative, and that high inflation rates increase uncertainty about future inflation. However, the relatively low and stable inflation rates in the West since the 1980s suggest that inflation risk premia have generally been small (and even negative) during this period, consistent with estimates in Grishchenko and Huang (2012).

2.3.3 Liquidity

The liquidity risk premium for a security is the compensation market participants require for taking on the risk of not being able to get out of a position in a timely manner to a reasonable price. A liquid market can be characterized by one in which investors are able to trade large volumes quickly, without moving prices too much.

2.3.4 Volatility

The duration of a bond does not tell the entire story of how large the effective interest rate risk is for that bond; we also need to consider the interest rate volatility. If, for example, a certain interest rate that is related to a bond A with a modified duration of 3% has a much higher standard deviation than the interest rate on bond B, that has a modified duration of 4%, then bond B might still be safer than bond A with regard to effective interest rate risk.

3 Previous Research

Due to their short history, most research on inflation-indexed bonds is relatively new. Since the inception of TIPS in the US in the late 1990s, some papers have been written on the predictability in the spread between nominal and inflation-indexed bonds. The aforementioned evidence of foreseeable variation in expected excess return of US government bonds stands in stark contrast to the expectation hypothesis. Pflueger and Viceira (2011) analyze the expectations hypothesis empirically for not only nominal bonds, but also for inflation-indexed bonds. In both US and UK markets it is rejected on both accounts, thereby further solidifying the existence of predictability in returns. Further, they investigate the spread between nominal and real bond returns, known as the break-even inflation risk premium, and are able to substantiate that it varies over time. Pflueger and Viceira make a case that the time variation in inflation-indexed bonds most likely indicates both a fluctuation liquidity risk premium, and a shifting real interest rate risk premium.

D'Amico, Kim and Wei (2010) use an affine term structure model to provide evidence for a quite large liquidity premium in TIPS. They estimate their model with the help of nominal and TIPS yields, inflation rates, and survey forecasts of interest rates. Estimates of inflation expectations and inflation risk premia that do not take the liquidity of TIPS into account become misleading, and produces large pricing errors for TIPS. However, when incorporating a liquidity factor into their model, a much better fit for these variables is gained. They obtain reasonable inflation risk premia and TIPS prices from the model and their estimate of inflation expectations correspond well to survey forecasts of inflation. In their estimation, the liquidity premium was approximately 1-2% in the earlier years of TIPS, but has come down notably in recent years. Their results indicate that liquidity-adjusted break-even inflation can provide useful information about real rates, expected inflation and inflation risk premia.

Grishchenko and Huang (2012) estimate inflation risk premia using TIPS data over the period 2000-2008, applying an approach that is arbitrage-free and largely model-free. They consider three different measures of TIPS liquidity, and estimate the liquidity premium to be around 13 bps over the full sample, but substantially higher in the beginning of the period. Depending on the proxy used for expected inflation, the average 10 year inflation risk premium they find varies from -9 to 4 bps over the entire sample.

Pflueger and Viceira (2013) estimate the difference in liquidity between nominal and inflation-indexed bonds to test for time-varying real interest rate, inflation and liquidity risk premia. They find strong and model-independent evidence that real interest rate and inflation risk premia contribute to the predictability in excess returns for nominal bonds. The estimated differential liquidity risk premia are systematic and range between 30 bps in 2005 to over 150 bps during the most recent financial crisis. They also find that liquidity risk seems to contribute to return predictability in inflation-indexed bonds. Furthermore, the authors examine and distinguish between three sources of excess return predictability in inflation-indexed bonds in the UK and the US, namely, liquidity risk, real interest rate risk and market segmentation. Additionally, they find that inflation risk premia exhibit significant time-variation, and are low on average, but take both positive and negative values in their sample. Strong empirical evidence is presented for time-varying real interest rate and liquidity risk premia in inflation-indexed bonds in general, and the liquidity risk premium is estimated to be the source of almost one-fifth of total TIPS excess return over the period.

Similar to our approach, some studies have analyzed the term structure of break-even inflation. Using a model of real and nominal yields, Adrian and Wu (2010) estimate the term structure of expected inflation. They document the discrepancy, as it is given by their model, between break-even inflation and inflation expectations (especially in times of high volatility in the market). Ejsing, Garcia and Werner (2007) estimate and use the term structure of break-even inflation rates for the euro area to obtain information about inflation expectations. Hördahl (2008) utilizes a term structure model to specifically study the inflation risk premium in both the US market and the euro area. A positive, but small and time-varying, wedge between break-even inflation and the expectations for future inflation of investors is identified.

4 Data

4.1 Yield Data

It is quite challenging to construct a proper term structure of interest rates, especially for real rates. Gaps in maturities of available bonds make interpolation necessary, and things like preferential tax treatments and liquidity premiums on certain bonds will affect yields. Real bonds also have the added complexity of indexation lags. Coupled with the fact that bonds have different coupon rates (affecting durations), it is not simple to find term structure data of high quality.

We use nominal Treasury and TIPS yields from authors Gürkaynak, Sack and Wright (2006, 2008). They construct smoothed zero-coupon yield curves for nominal Treasuries from January 1961 to the present, and for TIPS from January 1999 to the present. The yield curve is fitted to off-the-run nominal Treasuries and all TIPS issues[8] according to the Nelson-Siegel-Svensson model:

$$y(\tau) = \beta_0 + \beta_1 \left[\frac{1 - \exp\left(-{}^{\tau}/_{\lambda_1}\right)}{{}^{\tau}/_{\lambda_1}} \right] + \beta_2 \left[\frac{1 - \exp\left(-{}^{\tau}/_{\lambda}\right)}{{}^{\tau}/_{\lambda_1}} - \exp\left(-{}^{\tau}/_{\lambda_1}\right) \right]$$
$$+ \beta_3 \left[\frac{1 - \exp\left(-{}^{\tau}/_{\lambda_2}\right)}{{}^{\tau}/_{\lambda_2}} - \exp\left(-{}^{\tau}/_{\lambda_2}\right) \right]$$

where $y(\tau)$ is the yield at maturity τ, and β_0, β_1, β_2, β_3, λ_1 and λ_2 are parameters to be estimated.

By fitting a term structure to the data using a parametric model we do not get the best possible fit, but the model should be adequate for the purpose of analyzing underlying fundamental market risk premia. Since we do not concern ourselves with potential idiosyncratic effects at specific maturities, stemming from for example increased demand for a cheapest to deliver bond at bond futures settlement, this data should be satisfactory. As a result of this method of fitting the data, various idiosyncrasies at particular maturities are often missed, but the fit overall is good. Gürkaynak, Sack and Wright examine the differences between their fitted curves and observed yields until 2007. The largest deviations are at the shortest (2-5 year) and longest (20-30 year) maturities. The deviations are usually quite negligible, only a few basis points,

[8] As noted by Gürkaynak, Sack and Wright (2008) there is no liquidity premium in on-the-run TIPS. If there had been a premium though, it would have affected the fitted yields, and therefore also our results.

but are significant for the years 2002-2003 when the deviations peaked at around 25 bps for the short-term rates.

TIPS with maturities under 5 years will be excluded in our analysis due to the sensitivity of such short-term bonds to the indexation lag, and to short-term outcomes and seasonal effects of the CPI-U. TIPS with maturities longer than 20 years are also excluded due to considerations relating to the convexity of such long-term bonds (resulting in poorer fit).

Table 1

Descriptive Statistics for Break-even Inflation

The following table provides summary statistics for the yield data used in this study. The time period is from January 1999 to January 2014 at monthly intervals. Break-even inflation is abbreviated "BEI". Even though the series of 2-4 year maturities are excluded from any further analysis, they have been included here in order to facilitate comparison. I(O) represents the order of integration. A stationary series is I(0).

Series	Mean	Sd	Variance	Skewness	Kurtosis	I(O)	N	Min	Max
BEI02	1.66	1.19	1.42	-2.40	11.19	0	121	-4.24	3.19
BEI03	1.76	0.99	0.99	-2.42	11.05	0	121	-3.18	3.05
BEI04	1.86	0.84	0.71	-2.46	11.05	0	121	-2.37	2.94
BEI05	1.86	0.64	0.41	-2.14	10.86	0	181	-1.69	2.85
BEI06	1.95	0.56	0.32	-2.04	10.13	0	181	-1.11	2.79
BEI07	2.03	0.50	0.25	-1.96	9.64	0	181	-0.61	2.76
BEI08	2.11	0.44	0.20	-1.89	9.23	0	181	-0.19	2.76
BEI09	2.19	0.40	0.16	-1.81	8.86	0	181	0.16	2.80
BEI10	2.26	0.37	0.14	-1.73	8.49	0	181	0.39	2.87
BEI11	2.31	0.35	0.12	-1.63	8.13	0	181	0.56	2.93
BEI12	2.36	0.33	0.11	-1.54	7.77	0	181	0.69	2.98
BEI13	2.41	0.32	0.11	-1.45	7.45	0	181	0.78	3.03
BEI14	2.44	0.32	0.10	-1.38	7.16	0	181	0.85	3.08
BEI15	2.47	0.32	0.10	-1.31	6.94	0	181	0.90	3.12
BEI16	2.49	0.33	0.11	-1.27	6.77	0	181	0.92	3.15
BEI17	2.50	0.33	0.11	-1.24	6.68	0	181	0.93	3.18
BEI18	2.51	0.34	0.12	-1.23	6.64	0	181	0.92	3.21
BEI19	2.52	0.35	0.12	-1.24	6.67	0	181	0.90	3.23
BEI20	2.52	0.36	0.13	-1.26	6.75	0	181	0.88	3.25

4.2 Data on inflation expectation proxies

From Bloomberg, we obtain data on long-term (5-10 year) inflation forecasts compiled by the University of Michigan. Furthermore, we make use of the Chicago Fed National Activity Index

(CFNAI), which is a monthly weighted average of 85 indicators of US economic activity, as our proxy for short-term (1 year) inflation.

4.3 Data on liquidity proxies

We use data from Gürkaynak, Sack and Wright for our off-the-run 10-year par yields, and from Bloomberg, we use the USGG10YR as our on-the-run 10-year nominal yields. The difference between these series will represent our off-the-run spread variable (*OFFTR*).

Daily data on asset swap Z-spreads are acquired from Credit Suisse First Boston (CFSB).[9] However, data on the asset swap Z-spreads only go back to June 2003 (limited by the short history of asset-swapped TIPS). Therefore, in the same fashion as Pflueger and Viceira (2013), we equate those missing values to that of June 2003. The differences between the Z-spreads of asset-swapped TIPS and nominal Treasuries are used to create the relative Z-spread variable (*RZS*).

The Federal Reserve Bank of New York offers data on the transaction volumes of inflation-indexed and nominal Treasuries.[10] This data is used to create the relative transaction volume variable (*RTV*), but also in the creation of the turnover ratio (*TOR*) for inflation-indexed debt. For data on the outstanding amount of TIPS every month, we have made use of the Monthly Statement of Public Debt published by TreasuryDirect.

4.4 Data on volatility and inflation risk proxies

We use the Merrill Lynch Option Volatility Estimate Index (MOVE), which we acquired from Bloomberg, as our volatility proxy. For our inflation risk proxy we use seasonally-adjusted and monthly CPI-U data from the St. Louis Fed's Federal Reserve Economic Data (FRED) database. To create monthly volatilities of the CPI-U calculated over a rolling 12-month window, we use data beginning in 1998.

---

[9] For our daily data on yields and asset swap Z-spreads, we select the last observation of a month and let his value represent that month. Theory on our proxy variables and more information on how they are created is presented in section 5.

[10] Primary dealer transactions reported to the Federal Reserve Bank of New York. Due to changes in the data structure, the data is split into three different time-periods. We use transaction data for longer maturities; pre-July 2001 we use the 5MOT (nominal yields, maturities > 5 years) and TIIS (Treasury Inflation-Indexed Security) data. Pre-April 2013 we use the 611OT (nominal, 6 years < maturities <=11 years), and TIIS. Finally, for April 2013 and later, we employ the G7L11 (nominal, 7 years < maturities <= 11 years), and G6L11 (TIPS, 6 years < maturities <= 11 years).

Table 2

Descriptive Statistics for Regressors

The following table provides summary statistics for the variables used as regressors in this study. The time period is from January 1999 to January 2014 at monthly intervals. MFI is the University of Michigan survey on forecasted long-term (5-10 year) inflation, CFNAI is the Chicago Fed National Activity Index, OFFTR is the off-the-run spread, RZS is the relative Z-spread (in basis points), RTV is the relative transaction volume (log of the ratio of TIPS to nominal bond transaction volumes), TOR is the turnover ratio of TIPS (ratio of transaction volume to amount outstanding, in dollar terms), MOVE is the Merrill Lynch Option Volatility Estimate Index, and CSD is a series of monthly standard deviations of the CPI-U. I(O) represents the order of integration. A stationary series is I(0).

Series	Mean	Sd	Variance	Skewness	Kurtosis	I(O)	N	Min	Max
MFI	2.88	0.14	0.02	0.65	4.05	0	181	2.50	3.40
CFNAI	-0.29	0.90	0.80	-2.01	8.75	0	181	-4.52	1.12
OFFTR	0.18	0.12	0.02	0.86	3.85	1	181	-0.02	0.64
RZS	34.40	17.78	315.96	3.03	14.15	1	181	15.09	134.56
RTV	-1.34	0.23	0.05	-0.87	2.65	1	181	-1.95	-1.04
TOR	0.07	0.03	0.00	0.96	3.58	1	181	0.01	0.18
MOVE	21.61	7.62	58.13	1.30	5.54	1	181	10.96	51.72
CSD	5.76	1.98	3.92	0.48	2.18	0	181	2.76	9.76

5 Methodology

5.1 Ordinary Least Squares (OLS) Regression

We analyze the term structure of break-even inflation by using principal component analysis (PCA).[11] In order to determine which risk factors explain the principal components and the 10-year break-even inflation (for comparison with other studies), we regress these on our risk proxies using the following (Equation (1)):

$$\begin{pmatrix} PC1_t \\ PC2_t \\ PC3_t \\ BEI10_t \end{pmatrix} = (\beta_1\ \ \beta_2)\begin{pmatrix} MFI_t \\ CFNAI_t \end{pmatrix} + (\beta_3\ \ \beta_4\ \ \beta_5\ \ \beta_6)\begin{pmatrix} OFFTR_t \\ RZS_t \\ RTV_t \\ TOR_t \end{pmatrix} + (\beta_7\ \ \beta_8)\begin{pmatrix} MOVE_t \\ CSD_t \end{pmatrix} + \varepsilon_t$$

where

t = 1, 2, ..., 181 (monthly data)

PC stands for principal component.

BEI10 is the 10-year break-even inflation series.

MFI is the Michigan inflation forecast.

CFNAI is the Chicago Fed National Activity Index.

OFFTR is the off-the-run spread.

RZS is the relative Z-spread.

RTV is the relative transaction volume.

TOR is the turnover ratio.

MOVE is the Merrill Lynch Option Volatility Estimate Index.

CSD is the monthly CPI-U standard deviation (rolling 12-months, annualized).

ε is a residual component of the model with an expected value of zero.

The OLS regressions are performed using robust standard errors since the Breusch-Pagan/Cook-Weisberg tests for heteroskedasticity confirm that we do not have constant variance of the residual (see Table 8 in Appendix A).[12] Because the variables take on such different values (see

[11] See Appendix B for theory on principal component analysis.

[12] The Gauss-Markov Theorem, which states that OLS is the best linear unbiased estimator (BLUE), relies on the assumption of homoskedasticity, i.e. that the variance of the unobserved error conditional on the independent variables $Var(u|x)$, is constant. However, heteroskedasticity does not cause bias nor inconsistency in the OLS estimators, it simply means that the usual *t* statistics are no longer valid since the OLS standard errors used to create these are based on the estimators of the variances $Var(\widehat{\beta_j})$. The most straightforward way to tackle heteroskedasticity in a linear probability model is to continue using OLS estimation with the addition of robust standard errors in test statistics (Wooldridge, 2013).

Table 2 for descriptive statistics for the regressors), we standardize both the dependent and independent variables to a mean of 0 and a standard deviation of 1 before running the robust regressions. This will help in the interpretation of the regression results.

In the following subsections we will describe the variables used in Equation (1) above, e.g. how they are formed and what aspects of the theoretical risk premia of break-even inflation they are meant to capture.

5.2 Inflation expectation proxies used in the regression analysis

In an effort to better separate any liquidity effects that may exist in the principal components, or in the 10-year break-even inflation series, we include a proxy for long-term inflation expectations. Other studies such as D'Amico, Kim and Wei (2010), Grishchenko and Huang (2012), and Pflueger and Viceira (2013) have proposed this methodology.

In their in-depth paper, Ang, Bekaert and Wei (2006) find that, out of the four main methods used for forecasting inflation, survey forecasts outperform the other three approaches.[13] They analyze three different surveys. The Livingston survey and the Survey of Professional Forecasters (SPF) are conducted amongst professionals (the former asks economists from academia, government and industry whereas the latter uses estimates primarily from business). The survey from the University of Michigan, on the other hand, asks households, i.e. consumers, for their estimates of the expected price changes in the future.

Even though the professionals consistently performed better inflation forecasts than the consumers of the Michigan survey, we believe that this survey would be the one most suited to employ in our analysis. First of all, Ang, Bekaert and Wei point out that the forecasts of the Michigan survey are only slightly worse than those of the Livingston and SPF surveys (and the forecasts are on a comparative level). Secondly, whereas the Michigan survey is conducted monthly, the SPF and Livingston surveys are carried out quarterly and semi-annually, respectively. For Ang, Bekaert and Wei, who use data going back as far as 1952 in the case of the Livingston survey, the frequency with which the surveys are conducted is most probably not of any greater concern. For our purposes however, with a time span from January 1999 to January 2014, large gaps in data could result in interpolations (in order to get forecasts of inflation on a monthly basis) that fail to account for crucial variance of our dependent variables. Therefore, we

[13] The other methods are: forecasts from the yield curve, time-series forecasts and forecasts based on the Phillips curve. For readers interested in an elaboration on these and how they perform, we recommended Ang, Bekaert and Wei (2006).

make use of the Michigan survey in our analysis to proxy for US inflation expectations.[14] Considering the maturities of the debt instruments used in this study (5-20 years), we use the longer-term inflation forecasts over the horizon of 5 to 10 years (instead of the alternative which is a one year horizon).

Long-term inflation forecasts should not contain indications of shorter-term inflation expectations, but there is, however, a possibility that break-even inflation contains information about expected inflation in the near future. A proposition of Pflueger and Viceira (2013) is therefore to utilize the Chicago Fed National Activity Index (CFNAI). The idea behind this index of economic activity was developed by Stock and Watson (1999) who restrict their attention to looking at forecasts of US inflation at 12 month horizons. They consider the possibility that for the different methods of gauging future inflation, there exist some factor that is shared between these indicators, and for predicting inflation it is this factor, or index, which is really of use. Stock and Watson (2002, 2003) provide further evidence that combining indices (and a possible averaging of these) improves the forecasting ability of macroeconomic series. They find that by averaging individually unreliable forecasting series, a reliable combination forecast can be generated.

The CFNAI is a weighted average of 85 monthly indicators of national economic activity for the US. These economic indicators are drawn from four broad categories of data: (1) production and income, (2) employment, unemployment, and hours, (3) personal consumption and housing, and (4) sales, orders and inventories. Consequently, by using CFNAI, we have in a single measure a common factor from many aspects of the general macroeconomic activity. The idea is for this index to adequately account for short-term inflation.[15]

5.3 Liquidity proxies used in the regression analysis

As highlighted by many researchers, there was a large difference between the liquidity of inflation-indexed bonds and their nominal counterparts for many years following the inception of TIPS in the US (see Elsasser and Sack (2004), Adrian and Wu (2010); among others). Some papers further describe and show that a significant improvement of these conditions occurred around 2003, see for example Elsasser and Sack or Grishchenko and Huang (2012). Instead of

[14] See Figure 4 in Appendix for a time-series comparison of the Michigan surveys (1 and 5-10 years) and SPF surveys (1 and 10 year). Data on the SPF surveys have been collected from Federal Reserve Bank of Philadelphia. The SPF surveys are conducted in the middle of each quarter and we have interpolated them to a monthly basis (for this we used data from the last quarter of 1998 to the first quarter of 2014). The long-term expected inflations (both SPF and Michigan) are not surprisingly the most stable whereas the short-term SPF, and especially the 1-year Michigan, is very volatile.

[15] The correlation between CFNAI and the Michigan 5-10 year survey forecasts is -0.10.

excluding the years before this illiquidity was overcome, we, like Adrian and Wu, include the early years and make use of variables which we hope will be able to account for the relative differences in liquidity.

In his model, which applies to the OTC market, Weill (2007) emphasizes the search frictions that arise in bilateral trade. Buyers and sellers need to allocate time to find one another, meet, and then negotiate the price. Therefore, the more liquidity an asset is associated with, the quicker buyers and sellers for that asset can be found, and thus, the smaller the trading delays will be. The ease of buying an asset, and selling it, is linked to the amount of tradeable shares of said asset. As tradeable shares increase, larger trading volumes and turnover are implied, and Weill points out that we can proxy for liquidity using the latter two of these three measures.

Furthermore, the results of Fleming and Krishnan (2012) substantiate that trading activity would likely be a better proxy for liquidity cross-sectionally for TIPS, i.e. across different maturities, than bid-ask spreads. The difference between bid and offer prices is a common measure of liquidity, see for example Fleming (2001), however, according to Fleming and Krishnan, this measure does not suit the TIPS market as well as it does the nominal bond market. Therefore, corresponding to the methodology of Pflueger and Viceira (2013), we use a relative measure of transaction volumes between TIPS and nominal bonds, which we call *RTV*. The monthly transaction volumes for the two types of securities are used to create a ratio.

$$MD_t = \log\left(\frac{TV_t^{TIPS}}{TV_t^{\$}}\right)$$

where MD_t is the log of the ratio of transaction volumes (*TV*) at month *t*. Superscripts \$ and TIPS are used to indicate nominal securities and TIPS respectively, and following in the footsteps of Pflueger and Viceira, we smooth the data by creating a 3-month moving average:

$$RTV_t = \frac{MD_{t-2} + MD_{t-1} + MD_t}{3}$$

where RTV_t is the relative transaction volume at month *t*.

In addition to accounting for differences in the ease of finding counterparties and other search frictions in bond markets, we employ the off-the-run spread (*OFFTR*) as a liquidity proxy which we expect will capture the varying demand for Treasuries across time. The difference in demand between the latest issued securities ("on-the-run") and the older "off-the-run" is established in the literature. Longstaff (2002) provides evidence of a difference in yields between

Treasury and equivalent Refcorp bonds during "flights to liquidity", i.e. periods when market sentiment leads to a strong preference for holding the most liquid assets. Refcorp bonds are inherently guaranteed by the Treasury, and therefore they share credit risk (and the default-free status). Consequently, they only differ with regard to liquidity. Hence, we can expect that there have been periods when the aforementioned difference in liquidity between inflation-indexed and nominal bonds could have been temporarily widened by an increasing demand for on-the-run nominal bonds.[16]

$$OFFTR_t = y_t^{off-the-run} - y_t^{on-the-run}$$

where $OFFTR_t$ signifies the off-the-run spread of month t.

As discussed by D'Amico, Kim and Wei (2010), asset swap spreads can be used as measures of liquidity. Asset swap spreads are determined by what default and liquidity risks investors attach to the underlying asset. However, since inflation-indexed and nominal Treasuries are both regarded as having no risk of defaulting on payments, what remains as a determinant of spreads is essentially liquidity risk. D'Amico, Kim and Wei, as well as Grishchenko and Huang (2012), use the difference in ASWs for asset-swapped off-the-run and on-the-run nominal Treasuries to proxy for the relative illiquidity of inflation-indexed bonds. Yet we, again, emulate Pflueger and Viceira (2013) and use 10-year asset-swapped nominal bonds and TIPS.[17] Pflueger and Viceira argue that the difference between the asset swap spreads can be used to proxy for existing and expected future relative financing costs for TIPS and nominal Treasuries.

However, unlike Pflueger and Viceira who use par asset swaps, we employ asset swap Z-spreads. Choudhry (2005) states that when analyzing an asset swap, the traditional approach is to use the bond's yield-to-maturity when calculating the spread. Yet, for yield-to-maturity calculations there are some implicit assumptions[18] which cause problems when using this spread

[16] Fleming and Krishnan (2012) point out that there is a notable difference between on-the-run versus off-the-run TIPS when it comes to trading activity, where, much as in the nominal market, there is higher demand for on-the-run TIPS. However, this does not indicate that there is a liquidity premium present in on-the-run TIPS. They "merely" find that different measures of liquidity capture the variation of the data to a different degree, some better than others.

[17] D'Amico, Kim and Wei (2010) states that although they would have preferred to use the difference in asset swap spreads between TIPS and off-the-run nominals, they consider the history of TIPS asset swaps to be too short. They subsequently show that the difference between the latest-issued and older nominal securities is a sufficiently good proxy since the daily correlation between the two spreads (the one used and the preferred) is 0.90 between mid-March and November, 2006 and 2009 respectively.

[18] Choudhry (2006) discusses how the main disadvantage of the yield-to-maturity calculations is that it relies on the unrealistic assumption that each coupon payment is reinvested at the redemption yield which is thought of as remaining stable for the entire life on the bond. A further disadvantage of the yield-to-maturity measure is that it does not take into account the situation where investors do not hold the bond to its maturity, in which case the return on the bond will not be as great.

for relative analysis. Therefore, the Z-spread, which uses the zero-coupon yield curve in order to compute the asset swap spread, is considered more suitable. Hence, with the Z-spread, the cash flows from a bond are discounted using the particular zero-coupon rate that matches the maturity of that cash flow. The relative Z-spread could thereby be considered as the most "pure" measure of the asset swap. The difference to the conventional asset swap spread should be at its lowest for bonds with good credit-quality (such as those from the Treasury used in this study) and for shorter-term bonds. A large difference could be a sign of mispricing.

$$RZS_t = AZS_t^{TIPS} - AZS_t^{\$}$$

where RZS_t is the relative Z-spread at month t on 10-year asset swap Z-spreads (AZS). TIPS and $ superscripts denote TIPS and nominal spreads respectively.[19]

The turnover ratio is a measure of liquidity for inflation-indexed bonds only.[20] It is created by dividing the accumulated transaction volume of TIPS within month t with the amount of outstanding inflation-indexed debt at the end of the same month. We believe this measure might be able to capture a liquidity component inherent to the break-even inflation spread that our other proxies may not account for. While the total amount of outstanding marketable inflation-indexed debt over the period of January 1999 to January 2014 exhibits (almost) linear growth, the growth in nominal Treasury notes and bonds is much more moderate up until 2009 when it starts climbing sharply (see Figure 3). With an asymmetrical growth in the amount of debt outstanding for these two categories of instruments, the relative transaction volume, which has steadily improved over time until it somewhat stagnates in 2005 (see Figure 2), might not fully depict the improvement in liquidity of inflation-indexed debt over time. The relative transaction volume does not take into account that such a substantial increase in outstanding nominal debt in the latter half of our sample (compared to that of TIPS) should be accompanied by a comparable rise in trading volume. Therefore, while the Treasury issued much more debt not linked the general price level, a larger amount of inflation-indexed debt (relative to its outstanding amount) could actually have been trading.

$$TOR_t = \frac{\sum_{t=1}^{T} TIPSTV_t}{\sum_{t=1}^{T} TIPSOS_t}$$

[19] Our nominal data is on-the-run, however, this should not matter since if we used off-the-run nominal data for the asset swap spread it should only induce a difference in level. The relationship between the series should be the same.

[20] Unlike D'Amico, Kim and Wei (2010), and Sack and Elsasser (2004) whom use a 13-week moving average of the turnover ratio, we make this a monthly measure. We believe that by smoothing, we might somewhat eradicate important variation in the data.

where *TOR* is the turnover ratio, and *TIPSTV* and *TIPSOS* are the transaction volume and the outstanding amount of TIPS respectively.

As we can note in Figure 2, the turnover ratio was at its highest following the improvement in liquidity conditions around 2003, as earlier discussed, and peaked a few times at above 15%. However, due to the contemporaneous positive correlation between volatility in prices (and also asset returns) and trading volume, a relationship diligently researched and composed in Karpoff (1987), turnover might not be a measure without flaws since this correlation is found to be unrelated to the liquidity conditions of the market. Yet, with the small array of proxies for liquidity that have now been presented, we should adequately be able to pick up any liquidity component that is bound to be a part of the principal components and the 10-year break-even inflation.

The correlations between the liquidity proxies can be viewed in Table 9. The turnover ratio exhibits a very low correlation with the other measures, and the off-the-run spread, the relative Z-spread and the relative transaction volume are not highly correlated with each other.

5.4 Volatility and inflation risk measures used in the regression analysis

Fornari (2008) estimates volatility risk premia by studying the difference between implied volatilities from interest rate swaptions and the expectations of such swap rate volatilities. He finds that volatility risk premia are proportional to risk aversion and have been negative, strongly time-varying and exhibit a term structure. Compensation for volatility risk is positively related to expected volatility, and influenced by the level and volatility of the short-term rate. The major difference between Fornari's paper and our application is that we are studying the spread between nominal and real rates, and not just nominal rates. As far as we are aware, there is currently no research on volatility risk premia on real rate instruments.

We will simplify the analysis significantly in this paper and take a naïve approach to constructing a proxy for volatility risk. We use the *MOVE* (Merrill Lynch option volatility estimate) index outright to capture some of the characteristics of the time-variation in the relative volatility compensation between nominal bond and TIPS yields. The *MOVE* index consists of normalized (i.e. correcting for the price) implied volatilities of a wide variety of options with approximately 1 month to maturity on Treasury securities. The index is the weighted average of implied volatilities, weighted according to estimates of OTC volumes. The weights and underlying bond maturities of the index are 40% in options on 10-year bonds, and 20% in options of 2, 5, and 30-year bonds each. *MOVE* can be seen as reflecting market estimates of future short-term volatility of long-term Treasury yields.

Readers should note that there is a discrepancy between our theoretical construct of differential volatility risks as defined in section 2.3.1, which is what we ideally want to measure, and the *MOVE* index, which only measures the volatility of nominal bond yields. We would certainly agree that there most likely exist better proxies than *MOVE* for measuring the effect that relative volatility risk has on break-even inflation. However, as we will see in section 6, this proxy still does well based on our results, and it is such a parsimonious way to incorporate the volatility aspect into the analysis that it becomes a very suitable starting point for us.

Our proxy for the inflation risk premium is computed by taking the volatility over the rolling last 12-month period for the seasonally-adjusted CPI-U and annualizing it. We use the seasonally-adjusted series to avoid overestimating the impact temporary spikes would have on expectations. Since the seasonality of inflation is expected by market participants, it should not affect the risk measure.

$$CSD_t = \sqrt{12} \times CPIUSD_t$$

where CSD_t is the annualized CPI-U volatility at time t, and *CPIUSD* denotes the monthly standard deviation of the CPI-U computed over the past year. Figure 5 shows graphical representations of our volatility and inflation risk variables.

6 Results

6.1 Principal Component Analysis

Table 3

Principal Components

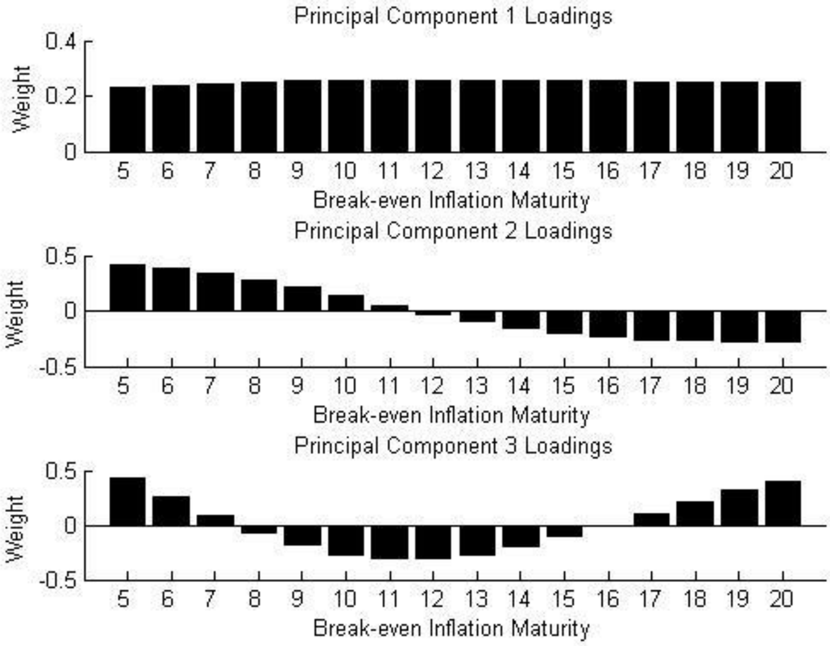
These principal components are the result of a principal component analysis which has been conducted for standardized (mean of 0 and a standard deviation of 1) break-even inflation series, 5-20 year maturities.

	Percent of the variance explained by each PC	Cumulative R^2 of including PC 1,2,3
Principal Component (PC) 1	93.23%	93.23%
Principal Component (PC) 2	5.73%	98.96%
Principal Component (PC) 3	0.99%	99.95%

We run a PCA on the term structure of break-even inflation rates for maturities from 5 to 20 years. As can be seen in Table 3 above, the first principal component (PC) explains over 93% of the variation in the data, and the first three PCs explain almost 100% combined. These results can be compared to those of Gürkaynak, Sack and Wright (2008) where the explained variance is more evenly distributed between the first and second PC, each explaining around 78% and 21% of the variance in the data respectively. The difference between their analysis and ours, apart from analyzing data for different time periods, is that they include the shortest maturity break-even inflation rates which we do not. This gives us a valuable hint that perhaps short-term outcomes of monthly CPI-U releases (affecting short-term rates) may explain some part of the time variation in PC2. We will look into this in more detail in section 6.3.

The factor loadings on the first three PCs seem to correspond to the three components traditionally used in the literature to describe the nominal yield curve, i.e. the level, slope and curvature (also called "shift", "twist", and "butterfly" respectively). See Figure 1 below for the factor loadings of these principal components. The first PC is basically equally weighted on all the different maturities. That can be interpreted as a parallel shift in break-even inflation rates across the entire term structure. The second PC is a slope component, i.e. a spread between the shorter and longer maturity break-even inflation rates. This PC explains the part of the data where the short-term bonds are positively correlated with the overall data and the long-term bond are negatively correlated with the general data. The third PC shows curvature and convexity characteristics. From the graphs of factor loadings, we can see why this principal component is often interpreted as a trough or hump in the term structure. In the Nelson-Siegel-Svensson model, the second and third factors can be seen as representing two humps that can be added to the specification of a yield curve.

Figure 1: Factor loadings of Principal Components 1-3



6.2 Regression Analysis Principal Component 1

For the first principal component (PC1), we can from Table 4 see that inflation expectations (both short- and long-term) describe 39% of the variation in the data (see column (1)). Coefficient estimates for both *CFNAI* and the Michigan 5-10 year survey are positive and statistically significant. From Figure 1 above, we know that PC1 loads similarly on all maturities of break-even inflation, and Table 10 (in Appendix A) provides further evidence of how closely PC1 resembles individual break-even inflation series.[21] The positive signs of the coefficients are consistent with theory and also make sense intuitively. If inflation expectations rise, one would expect that the spread between nominal and real yields would widen. Additionally, no sign alternates from positive to negative, or vice versa, as more explanatory variables are included in the regression.

[21] We can, from Table 10, see that PC1 is most correlated with the BEI maturities at the middle of the term structure. The resemblance of PC1 with 10-year break-even inflation can also be noted by looking at their graphs in Figures 6 and 7.

Table 4

Regression Results PC1

Regressions follow Equation (1) for the period from January 1999 to January 2014. Dependent and independent variables alike have been standardized to a mean of 0 and a standard deviation of 1. No constant is used - having standardized the variables makes any intercept redundant. The asterisks (*) after each coefficient signal the significance level of the result. *, ** and *** represents the 1, 5 and 10% significance levels respectively. Robust standard errors are stated in brackets below the coefficients. We incrementally add variables to the regression, starting off with the two measures for expected inflation in column (1). Columns (2) to (5) show the addition of liquidity measures. Columns (6) and (7) contain the further additions of the proxies for volatility and inflation risk premia respectively.

Factor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MFI	**0.23***	**0.18***	**0.18***	**0.18***	**0.18***	**0.19***	**0.13****
	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
CFNAI	**0.61***	**0.40***	**0.29***	**0.37***	**0.31***	**0.33***	**0.36***
	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
OFFTR		**-0.39***	**-0.37***	**-0.05**	**-0.18*****	**-0.31****	**-0.31****
		(0.06)	(0.06)	(0.11)	(0.11)	(0.12)	(0.12)
RZS			**-0.19***	**-0.42***	**-0.31***	**-0.32***	**-0.37***
			(0.07)	(0.09)	(0.09)	(0.10)	(0.09)
RTV				**0.39***	**0.22****	**0.20****	**0.16*****
				(0.09)	(0.10)	(0.10)	(0.10)
TOR					**0.22***	**0.21***	**0.19***
					(0.05)	(0.05)	(0.05)
MOVE						**0.17*****	**0.16*****
						(0.10)	(0.09)
CSD							**0.16****
							(0.07)
R^2	0.39	0.50	0.52	0.58	0.61	0.62	0.64

Our results indicate that for PC1, the *CFNAI* appears to be more economically significant than the long-term forecasts of the Michigan survey. This would suggest that although the long-term inflation forecasts of households in the US are able to contribute to the explanation of the variation in PC1, short-run inflation expectations captured by the *CFNAI* are even more important to include. This result is to be expected in light of Figure 4 where we can see that long-term inflation expectations have been remarkably well anchored around 3%. That is somewhat higher than the Federal Reserve's medium to longer term target of around 2% as measured by the core personal consumption expenditures (PCE) price index. Given the relatively modest inflation rates of the last two decades, and the Fed's high credibility, this have probably prevented short-term bouts of inflationary pressures from making their way into longer run expectations. Both short-term inflation expectations and break-even inflation rates have exhibited significantly more variation over time than their long-term counterparts as shown by their respective graphs in Appendix A (Figure 4 for inflation expectations and Figure 7 for 10yr BEI).

When adding our liquidity proxies to the right-hand side of the regression specification, R^2 increases to 61%. The liquidity element seems inherent to PC1, which is, as we should keep in mind, the single best representation of the variation across different maturities of break-even inflation. Although the most noteworthy contribution (at least in terms of increased R^2) comes from the off-the-run spread, Table 4 show how the three other measures, i.e. the relative Z-spread, the relative transaction volume, and the turnover ratio still appear to capture different aspects of the liquidity disparity.

The signs of the liquidity proxies are expected given how they were designed and what aspects of liquidity they represent. If the off-the-run spread increases, it indicates that there is a "flight to liquidity" to some degree in progress, showing that investors prefer on-the-run nominal securities that are more liquid. Since the coefficient is negative, the level of break-even inflation decreases because of the increased liquidity risk premium required to hold inflation-indexed securities relative to nominal securities. Neither does a negative sign for the relative Z-spread surprise considering that when this spread increases, it signifies a current or future broadening of the gap between financing costs for taking positions in nominal and inflation-indexed securities. Therefore, the off-the-run and the relative Z-spread could be regarded as embodying the illiquidity of the TIPS market. The relative transaction volume and turnover ratio, on the other hand, have both entered positively, which is also in agreement with how these variables are meant to pick up on the improving or worsening liquidity of TIPS relative to that of nominal Treasuries. When the relative transaction volume increases, the transaction volume of TIPS is moving higher compared to nominal Treasury volume, and as the turnover ratio rises, it displays how trading volume in inflation-indexed bonds as a fraction of bonds outstanding is improving. These are both consistent with break-even inflation increasing, due to the relative illiquidity of TIPS compared with nominal Treasuries.

A standard belief is that the gap between nominal and real yields should narrow as the differences in liquidity tighten. The gap is reduced if nominal yields decrease, or if real yields increase. However, there is no consensus in the literature whether episodes of significant narrowing of break-even inflation like in 2008 are due to one or the other. As mentioned briefly in section 5.3, some argue that a "flight to liquidity" in nominal Treasuries, resulting in a large premium for those bonds, should be seen as the main explanation. Others claim that it is rather a higher liquidity risk premium in TIPS that is the correct reason.

For PC1 we could lastly make a note of how the statistical significance overall was very good. In only one instance (column (4) Table 4) was there an estimate with a statistical significance above the 10% level. We see that the *MOVE* index is significant at the 10% level while our proxy for the inflation risk premium is significant at the 5% level. Additionally, the economic significances of these two proxies are estimated to be the same. The signs of the coefficients tell us that break-even inflation increases when yield volatility increases, which makes sense as nominal yields should be more volatile than TIPS yields under normal market conditions. Similarly, break-even inflation increases when our inflation risk premium proxy increases; an expected result since more inflation uncertainty should always lead to higher nominal bond yields ceteris paribus, while TIPS yields should remain unchanged by this.

6.3 Regression Analysis Principal Component 2

Table 5

Regression Results PC2

The table shows the results of Equation (1) with PC2 as the dependent variable. The content follows the outline of Table 4.

Factor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MFI	**0.32***	**0.27***	**0.27***	**0.29***	**0.29***	**0.26***	**0.14***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
CFNAI	**0.16***	**-0.04**	**-0.04**	**-0.20***	**-0.16****	**-0.23***	**-0.16****
	(0.06)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)
OFFTR		**-0.38***	**-0.38***	**-0.97***	**-0.90***	**-0.58***	**-0.58***
		(0.07)	(0.08)	(0.10)	(0.10)	(0.13)	(0.12)
RZS			**0.00**	**0.43***	**0.38***	**0.41***	**0.30***
			(0.08)	(0.08)	(0.09)	(0.09)	(0.09)
RTV				**-0.73***	**-0.64***	**-0.58***	**-0.67***
				(0.10)	(0.12)	(0.12)	(0.11)
TOR					**-0.11**	**-0.09**	**-0.14****
					(0.07)	(0.06)	(0.06)
MOVE						**-0.45***	**-0.49***
						(0.11)	(0.10)
CSD							**0.37***
							(0.08)
R^2	0.12	0.22	0.22	0.41	0.42	0.49	0.55

From the results presented in Table 5 we can deduce that inflation expectations matter very little for the second principal component of the term structure. Although the liquidity measures are able to enhance the R^2, suggesting that the variables in the regression model are somewhat able to explain the variation in PC2, their signs are at odds with those obtain for PC1 (except for that of *OFFTR*). Out of these variables, the relative Z-spread appears to offer the least amount of incremental explanatory power. In column (7), with the full Equation (1), we can notice that all the variables are statistically significant at a 5% level.

We should, however, bear in mind that while PC1 still resembled individual break-even inflation series (being the level factor), this does not hold true for PC2 (the slope factor).[22] Thus, where an economic interpretation of the estimates for PC1 could be connected to how the variables should affect the break-even inflation in a rather straightforward manner, such an approach is not applicable for the estimates concerning PC2.

We note that our inflation risk premium variable is highly significant here and contributes strongly to explain the variation in the data. That would be consistent with the popular view that the term premium consist largely of an inflation risk premium (which itself has a term structure). The liquidity variables are also of some importance here, and we will offer an interpretation of this result now. Assume that liquidity conditions deteriorate. The probability of a recession and lower inflation will then increase, as will the uncertainty about inflation, i.e. the inflation risk premium. This will lead to selling (dropping prices) and higher rates (increasing yields) at the front end of the TIPS yield curve relative to the long end, especially given the impact outcomes of the CPI-U has on the front end. Therefore, including bonds with maturities shorter than 5 years would further increase the contribution to R^2 of PC2 as indicated in section 6.1. This line of thinking is consistent with the estimated coefficient of the off-the-run spread variable, but unfortunately, the signs of the other liquidity variables do not fit the interpretation.

Given the comparably low impact that *CSD* had on PC1 compared to PC2 (both in terms of magnitude and increased R^2), we propose that PC2 could be regarded as taking account for an inflation risk premia.

6.4 Regression Analysis Principal Component 3

PC3 explains very little of the variation in the term structure of break-even inflation during our sample period. We have still chosen to include the results from our analysis of this component for completeness and because the component might matter more during certain, shorter periods

---

[22] Regressing the second principal component onto the 5-year break-even inflation series ($PC2_t = \beta * BEI05yr_t + \epsilon_t$) yields an R^2 of approximately 16%. See Table 10 for equivalent regressions on the 10, 15 and 20 year BEI-series, though no R^2 for these go above 7%.

of time. While the curvature component does not seem to matter much on average for longer periods of time (such as our sample period), it could very well be relevant as an explanation during a specific week of trading for instance. Hence, this component might still matter greatly to, for example, speculators whom might even be interested in intraday changes in the term structure of break-even inflation.

Controlling for short- and long-term expected inflation, we can note in Table 6 below that when including all the proxies for liquidity, the R^2 reaches 0.38. Further including the *MOVE* and *CSD* variable, the R^2 improves slightly. For the complete regression of Equation (1) (in column (7)), our long-term expected inflation measure along with our first two liquidity proxies become statistically insignificant. Any economical inference is again hard considering the nature of the third principal component. The weights of this curvature factor alternate from positive at our shortest maturities, to negative for the middle section of the spectrum, and back to positive again at the end of spectrum. Thereby it, like PC2, shares nothing of the resemblance with the break-even inflation series that PC1 does.[23]

Table 6

Regression Results PC3

The table shows the result on Equation (1) with PC3 as the dependent variable. The content follows the outline of Table 4.

Factor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MFI	**-0.12*****	**-0.12*****	**-0.10**	**-0.10**	**-0.10**	**-0.12*****	**-0.03**
	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.06)	(0.08)
CFNAI	**0.30***	**0.29***	**0.01**	**-0.05**	**-0.16**	**-0.20****	**-0.24****
	(0.09)	(0.09)	(0.11)	(0.11)	(0.11)	(0.10)	(0.10)
OFFTR		**-0.01**	**0.05**	**-0.17**	**-0.39***	**-0.19**	**-0.19**
		(0.11)	(0.08)	(0.13)	(0.15)	(0.19)	(0.16)
RZS			**-0.50***	**-0.34****	**-0.15**	**-0.13**	**-0.06**
			(0.13)	(0.15)	(0.15)	(0.14)	(0.15)
RTV				**-0.27****	**-0.55***	**-0.51***	**-0.45***
				(0.12)	(0.14)	(0.14)	(0.14)
TOR					**0.37***	**0.39***	**0.42***
					(0.07)	(0.06)	(0.08)
MOVE						**-0.27****	**-0.25****
						(0.11)	(0.12)
CSD							**-0.25***
							(0.08)
R^2	0.11	0.11	0.25	0.28	0.38	0.41	0.44

[23] As can be noted in Table 10, PC3 has an R^2 of 3% when regressed onto the 5-year BEI-series.

One possible interpretation of humps or troughs in the term structure is the following: Since the short end of the curve is controlled by central banks, and the long end is governed much by inflation expectations and the long-term potential growth in the economy, it is the belly of the curve (maturities around 10 years) that is used the most to express changes in expectations and the path for rates in the longer term.

We can see in Table 6 that the transaction volume measures are among the most statistically and economically significant variables. Notice that the signs of these two variables are opposite. When *RTV* goes down (nominal Treasury transaction volume goes up relative to TIPS transaction volume), PC3 increases, and the break-even inflation term structure is more likely to display a trough (see the PC3 factor loadings to easier understand this). This is an indication that certain Treasury collateral could be trading in special in the repo market, or is cheapest to deliver for futures settlement, causing temporary extra demand for these bonds at some maturity. However, this does not explain a trough spanning several maturities though, but rather, only point specific dips in the rates for certain maturities. An analogue argument would hold for the *TOR* variable if TIPS were on special in the repo market due to strong demand for those bonds. Yet, such a circumstance would have to be due to other reasons than for futures settlement since there is no futures market for TIPS.

6.5 Regression Analysis 10-year Break-even Inflation

The regression results for a particular break-even inflation rate on the curve (10 year break-even inflation in this case) are similar to the results from the analysis of the term structure (PC1) when it comes to R^2. This does not come as a surprise considering the aforementioned similarity between PC1 and the 10-year break-even inflation. The estimate for *CFNAI*, the short-term expected inflation proxy, is smaller by a miniscule amount (see Table 7 below, column (1)). This is in line with what would be anticipated considering how long-term inflation in the US tend to be relatively stable and short-term inflation do not have as much impact on the long end of the curve. Further, the estimates suggests that long-term forecasted inflation matter more for the 10-year BEI than for PC1, which could be connected to how this forecast essentially share the horizon of the former while the latter, however much it may resemble individual maturities, is a statistical mixture of the entire range of maturities used in this study.

Table 7

Regression Results BEI10yr

The table shows the result on Equation (1) with the 10-year break-even inflation as the dependent variable. The content follows the outline of Table 4.

Factor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MFI	**0.27***	**0.22***	**0.22***	**0.22***	**0.22***	**0.22***	**0.15***
	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
CFNAI	**0.59***	**0.35***	**0.28***	**0.35***	**0.30***	**0.32***	**0.36***
	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
OFFTR		**-0.44***	**-0.42***	**-0.16**	**-0.26****	**-0.37***	**-0.37***
		(0.06)	(0.06)	(0.11)	(0.11)	(0.13)	(0.13
RZS			**-0.13*****	**-0.32***	**-0.24****	**-0.25****	**-0.31***
			(0.07)	(0.09)	(0.10)	(0.10)	(0.09)
RTV				**0.32***	**0.20****	**0.18*****	**0.12**
				(0.09)	(0.10)	(0.10)	(0.10)
TOR					**0.16***	**0.15***	**0.12****
					(0.05)	(0.05)	(0.05)
MOVE						**0.15**	**0.13**
						(0.10)	(0.09)
CSD							**0.22***
							(0.07)
R^2	0.39	0.53	0.54	0.57	0.59	0.60	0.62

The first liquidity proxy, the off-the-run spread, is a bit more important in this case[24] – which is reasonable considering that this regressor is created using data on nominal par yields with the same maturity whereas PC1 loads rather equally across the term structure. All of the estimates for the liquidity measures carry the expected signs, where the reasoning goes as described in section 6.2. Upon adding the *MOVE* and *CSD* variables in column (7), the relative transaction volume is, however, no longer statistically significant at any conventional levels.

Using the estimated coefficients of Table 7 and the standard deviation of the 10-year break-even inflation (0.37) and the regressors', we can calculate each proxy's economic significance. We note that changes in the *OFFTR* closely followed by *CFNAI* seem to have the greatest impact on the 10-year break-even inflation in terms of a one standard deviation change in a risk proxy. This illustrates the economic importance of liquidity as well as of short-term expected inflation for the TIPS market. Our regression results indicate that an increase of 0.12 in *OFFTR* (i.e. a 1 standard deviation change) yields a 14 bps decrease (–0.37 * 0.37) in 10-year break-even inflation, all else equal. An increase in the relative Z-spread of 22 bps (1.24 standard

[24] Comparing the second columns of Table 4 and 7.

deviations) yields a similar effect. Interesting is also that a two standard deviation change in the inflation risk proxy (*CSD*) gives a meaningful increase of 16 bps in the break-even inflation. A one standard deviation increase in short-term inflation as measured by our proxy *CFNAI* is on average related with a 13 bps increase in the 10-year break-even inflation. It is quite noteworthy that changes in expected short-term inflation elicit such a sizable reaction. The appropriate interpretation here would probably not be that short-term inflation changes expectations of long-term inflation, or that 10-year break-even inflation is so sensitive to changes in short-term inflation. Instead, the inflation proxy should rather be seen as an indicator of short-term liquidity in this case. If we get an unexpected spike up in the CFNAI, it is a sign of abundant liquidity, which would explain why 10-year break-even inflation is predicted to rise in such a situation.

Lastly, we can compare our obtained results for the 10-year break-even inflation with those from Pflueger and Viceira (2013). Although we use a wider time horizon and another survey of long-term forecasted inflation (the Michigan survey whereas Pflueger and Viceira employ the 10-year SPF), our controls for short- and long-term expected inflation yield the same R^2. It could further be noted that in their study the off-the-run spread (specified in the same manner) is more able to explain the variation in the 10-year break-even inflation. Coupled with the fact that neither *MOVE* nor *CSD* considerably increases our R^2 when added, this results in Pflueger and Viceira attaining a slightly higher coefficient of determination even though their specified model do not include proxies for volatility and inflation risk.

6.6 Crisis Dummies for Liquidity

To control for the possibility that we are overestimating the impact of the liquidity risk premium we have included dummy variables for the financial crisis period that started in 2008. We hypothesize that investors demanded a higher compensation for holding these securities, especially TIPS with its documented lower liquidity, during such a period of high turmoil in the market. Out of our sample, the National Bureau of Economic Research (NBER) classifies the length of the recession to 18 months[25], which is a noteworthy part. In the Appendix (Table 11) we have included the results of a regression of the form in Equation (1) with the addition of dummy variables for the liquidity proxies, taking the value of 1 during this credit crisis (and zero otherwise).[26]

[25] The economy peaked in December 2007 and hit its trough in June 2009.

[26] For brevity we have not included the volatility risk proxy (*MOVE*) nor the inflation risk proxy (*CSD*). The regression thereby formally takes the form: $Y_t = \beta_1 MFI_t + \beta_2 CFNAI_t + \beta_3 OFFTR_t + \beta_4 RZS_t + \beta_5 RTV_t + \beta_6 TOR_t + \beta_7 OFFTR_t d_t + \beta_8 RZS_t d_t + \beta_9 RTV_t d_t + \beta_{10} TOR_t d_t + \epsilon_t$, where Y denote PC1, PC2, PC3 and BEI10, and d_t is the dummy variable for the crisis.

Consistent with the theory on "flight to liquidity", which, if ever, should be prevailing in times of unstable markets, the dummies for the off-the-run spread enter negatively – although only statistically significant at conventional levels for the third PC. Based on the significance levels, the only dummy that should be used to draw conclusions is that of the relative transaction volume. For the first principal component, and the 10-year break-even inflation, it depicts how during the crisis, this liquidity proxy takes on a substantially higher positive sign compared to the "regular" sign of the variable, i.e. not during the credit crisis. We regard this as indication that as the transaction volume of inflation-indexed bonds became more similar to that of their nominal equivalents during the crisis (TIPS become more liquid), their price would increase (yields decrease) and thereby the difference between nominal and inflation-indexed bonds yield (break-even) increases.

It seems like crisis dummies do not provide a particularly strong case for risk premia appearing only in times of significant market stress, but rather, only widen greatly from lower levels compared to normal times. This might tell us these risks are priced into yields continuously rather than being “switched on” and “switched off” during a crisis and during more normal market environments.

6.7 Correlation, Variance Inflation Factors and Serial Correlation

In order to validate the robustness of our results and investigate potential sources of problems that could be underlying when running the linear regressions, we begin by looking at the correlation of our explanatory variables (see Table 9 in Appendix A). As briefly discussed in section 5.3 earlier, the liquidity proxies are not highly correlated. The biggest cause of concern in this study when it comes to how interconnected the regressors are, is the relationship between the off-the-run spread and the MOVE. With a positive correlation of 0.77, they are undeniably very connected. As shown by the previous Tables 4 and 5 they do, however, work well together in these regressions and therefore seem able to capture different aspects of the variation in our first two principal components. Including *MOVE* (in the sixth columns) does not alter the sign of the estimated coefficient for *OFFTR*, but it changes the scale on the other hand. For the first PC, where *MOVE* enters positively, *OFFTR* takes on a more negative value, and vice versa for the second PC where the coefficient of *MOVE* is estimated with a negative sign. Conversely, for the last PC (Table 6) *MOVE* and *OFFTR* seem to capture the same aspects of variation as the inclusion of the former renders the latter statistically insignificant.

A solution to possible problems stemming from such a strong, positive relationship between *OFFTR* and *MOVE* would be to regress for example *MOVE* on *OFFTR* and thereafter include the latter along with the obtained residuals (substitute for the former) as variables in our regressions. These regressors will subsequently have zero correlation.[27]

Although correlations between predictor variables were of no discernible concern in all but a few cases, we follow up by performing a test for the presence of multicollinearity amongst our predictor variables. This is a statistical phenomenon of high correlation which may inflate standard errors thereby making the estimated coefficients statistically insignificant. Testing the collinearity of our explanatory variables, we find no evidence that raises alarm about multicollinearity. The variance inflation factors (VIFs), are all but one less than 5 (*OFFTR* has the highest value at 5.35) with a mean (uncentered) VIF of 2.81.[28]According to Chatterjee and Hadi (2012), an established rule of thumb states that evidence of multicollinearity exists if: (1) The largest VIF is higher than 10, and (2) the mean of all the VIFs is substantially larger than 1. O'Brien (2007) does not discuss mean VIF values, but mentions that a value of 10 is a generally established limit for what should advocate measures to be taken (such as dropping or combining variables into an index). Although our average VIF is larger 1, no value even remotely approaches the double digits. Therefore, we have not found any indication that prompts for the exclusion of a certain regressor from our analysis.

Returning to the Gauss-Markov theorem, OLS requires both homoskedasticity and serially uncorrelated errors in order to be BLUE. With the presence of positive serial correlation in errors, the standard OLS variance formula[29] understates the true variance of the OLS estimator. This, in turn, leads to the belief that the OLS slope estimator is more accurate than what it actually is (Wooldridge, 2013).

When testing for the presence of autocorrelation in errors[30] using the Durbin-Watson D-statistic, Durbin's alternative test and the Breusch-Godfrey test for higher-order serial correlation, these show that we can reject the null hypothesis of no serial correlation (see Table 13). While the

[27] For brevity, results from such an implementation are excluded, however, it can be noted that this approach does not alter the coefficients of determination (nor estimates or standard errors of variables other than *OFFTR* and the *MOVE*-residuals). With this method the estimates on these two variables for PC1 have their magnitudes somewhat reduced whereas for PC2 *OFFTR* takes on an even more negative value (thereby becoming the most economically significant measure by far). For PC3, we notice a change where both variables are able to contribute with statistical significance at conventional levels (with negatively estimated coefficients). However, upon replacing *MOVE* with *MOVE*-residuals for the 10-year break-even inflation, we still find no contribution.

[28] See Table 12 in Appendix A for uncentered variance inflation factors. One can either calculate centered or uncentered VIFs. As noted by Groβ (2003), it only makes sense to calculate the centered VIF if you have a constant term included, which we do not in this study. Therefore, we use the uncentered VIFs. The uncentered VIF for variable x_i is calculated as $VIF_{uc}(x_i) = \frac{1}{1-\hat{R}_i^2}$ where $\hat{R}_i^2$ indicates the squared uncentered multiple correlation coefficient from regressing x_i without an intercept on all the other independent variables (where the constant term is included), see Belsley (1991).

[29] $Var(\widehat{\beta_1}) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{sigma^2}{SST_x}$

[30] These residuals can be viewed in Figure 8.

first test requires regressors to be strictly exogenous, the latter two does not. A standard approach when attempting to deal with serial correlation in an OLS regression is to incorporate a lagged dependent variable, according to Keele and Kelly (2005). Table 14 provides results for when the model has been changed so that either one or two lags of the regressand is included as regressors. We can note how Durbin's alternative test and the Breusch-Godfrey test still show that there is evidence of serial correlation for PC1 and for the 10-year break-even inflation. [31] For PC2 one lag is sufficient while PC3 requires two lags to remove the persistence from the error terms.

To qualitatively assess the autocorrelation, we investigate the degree of decay in the autocorrelation of the regression residuals. We can from the autocorrelation functions (ACFs) displayed in Figure 10, deduce that overall we do not seem to have persistent serial correlation. On the whole, the ACF values diminish quickly and stay within the tolerance bounds after a few lags.

[31] Since the Durbin-Watson D-statistic requires the regressors to be strictly exogenous, this test cannot be applied to models with lagged dependent variables since the error term will not be uncorrelated with the regressors.

7 Conclusion

The purpose of this study has been to investigate whether the variation in the term structure of US break-even inflation, defined as the difference between nominal and real yields, can be explained by proxies for liquidity, volatility and inflation risk as well as expected inflation. Our results indicate that this is the case.

Using the technique of principal component analysis, we find that there are three principal components that can account for virtually all variation in the 5-20 year term structure of break-even inflation. The first principal component, a variable equally weighted on all different maturities, explains 93% of the variation in the maturity spectrum. Our collection of risk proxies can explain this principal component rather well, yielding an R^2 of 64% for our full regression model. An especially good variable to indicate the liquidity component of the first principal component is the relative Z-spread – included to capture differences in current and expected future financing costs between nominal and inflation-indexed securities.

Our inflation risk proxy is highly statistically significant for all three principal components and contributes meaningfully to explaining time variation in particularly the second principal component. This could reflect an upward sloping term structure in inflation risk premia in nominal bonds, consistent with previous research done on this topic. The liquidity variables are also relevant for explaining the second principal component, probably because tightening liquidity has a large effect on short-term TIPS yields through decreases in CPI inflation.

In order to facilitate comparison with previous research and to study how our proxies for different risks can explain a single series, we have utilized the 10-year break-even inflation. In addition to the off-the-run spread and the relative Z-spread, short-term expected inflation and the proxy for the inflation risk premium are the most statistically and economically significant explanatory variables in this case. As an example, a 20 bps (approximately 1 standard deviation) increase in the relative Z-spread results in a 13 bps decrease in the 10-year break-even inflation.

While we believe that it would have been even more preferential to include a measure that accounts for differential volatility risk, we have settled for an index which accounts for only the volatility risk in nominal securities. For the first principal component it indicates how an increase in the volatility of nominal yields would drive break-even inflation upwards. We consider the overall good performance of this measure (which theoretically only captures one side of what we were looking for) as an indication that it could be worthwhile to substitute this proxy for one more suited to represent relative volatility risk.

The Merrill Option Volatility Estimate and the volatility of the CPI-U, our last two explanatory variables, are more than anything, a first crude attempt to incorporate proxies for these risks. More elaborate approaches like using ARMA modeling of inflation measures to study the residuals would probably be more successful. Using better data like implied volatilities of swaptions, and comparing them to fitted GARCH models would likely mimic volatility risk better. Trolle and Schwartz (2010) construct swap rate distributions from swaptions to study the variation in conditional moments of the distributions. They find that certain higher moments of market participants' "belief distributions" for macroeconomic variables can explain swap rate volatility, volatility risk premia (defined as the differences between physical and risk-neutral volatility), skewness, and skewness risk premia. Their regression results show that GDP beliefs and inflation beliefs are the most important explanatory variables for the USD and EUR market respectively.

On the whole, our most compelling finding is that no matter if one is trying to explain the variation in a single series of break-even inflation or across a wide spectrum of maturities, once short- and long-term inflation expectations are controlled for, it is crucial to account for the liquidity factor. The presence of such a systematic and economically significant liquidity component in break-even inflation implies the need for caution when using this spread as a measure of market inflation expectations.

Some concerns about correlation between our regressors have been raised and especially the strong, positive relationship between the off-the-run spread and the Merrill option volatility estimate has been evaluated. Although these two co-move to a high degree – which is not surprising considering how a higher (lower) instability should increase (decrease) the difference in yields as on-the-run securities become more (less) popular – we find that our obtained estimates can still be used to draw inference. Furthermore, no evidence of multicollinearity between our explanatory variables is found.

We have extended our robustness checks by testing for serial correlation in errors (a plausible drawback which we believe is ignored on a regular basis) in order to thoroughly assess how valid our results truly are. While we find convincing evidence that such an autocorrelation in the regression residuals does exist, it is, however, not persistent, i.e. it rapidly dissipates within a few lags. The presence of serial correlation might warrant the specification of another model as opposed to our use of the ordinary least square (OLS) regression (with robust standard errors to combat heteroskedasticity). A list of such, more appropriate models in the face of serially correlated error terms include (but does not limit itself to) generalized least square (GLS) estimators, feasible GLS, and one can also use differencing to deal with highly persistent data (Wooldridge, 2013). However, such a specification lies outside the scope of this thesis.

Another important observation is that our dependent variables are stationary, while some of our independent variables have a unit root (see Table 15 in Appendix A). This may well be a source of the serial correlation in our regression residuals, as it is hardly optimal to try to explain the behavior of a stationary variable with non-stationary variables, especially if those non-stationary variables are not cointegrated. We have not used first differencing of variables in our regression for the reason that interpretation of the results becomes difficult.[32] Hence, we have a situation where there is a trade-off between running the risk of getting spurious regression results and interpretability of results, and we decided to prioritize the latter.

[32] Tentative results from applying a first difference on regressands and regressors alike indicate that although this effectively addresses autocorrelation, it impacts the explanatory power of the model.

8 Further research

As shown by Rezende (2013), macroeconomic fundamentals contain valuable information about the risk premia in bonds. He utilizes half a dozen variables with close connection to the US business cycle in order to extract macro risk factors. These factors are then used to predict the excess returns of zero-coupon Treasury bonds. This provides evidence that when researching the predictability of bond risk premia, not only the yield curve should be considered – macroeconomic factors offer explanatory power as well. Therefore, we recognize that such an addition could possess some value added. However, given the amount of explanatory variables included in the full model (eight regressors), it might be problematic adding new ones as their variation might coincide heavily with an already specified inclusion. That being said, space for new additions such as a broad economic index could be made by dropping already established ones. But, such a decision is not to be taken hastily considering that the current predictor variables are employed to capture different aspects of risk (and expected inflation). Though, given the amount of liquidity proxies included (four), a possible exclusion would probably be one of these (most likely either the off-the-run spread since it correlates so highly with the volatility index or the turnover ratio considering that this is not a relative measure).

Since we have not incorporated any tax or macroeconomic variables in our analysis, the Economic Policy Uncertainty Index could be an interesting and parsimonious addition to the model[33]. This measure of policy-related economic uncertainty consists of three types of underlying components. One component quantifies newspaper coverage of policy-related economic uncertainty, a second component reflects the number of federal tax code provisions set to expire in future years, and the third component uses disagreement among economic forecasters as a proxy for uncertainty.

Perhaps the most obvious ways to improve our model would be to find better volatility and inflation risk proxies. Fitting a model for volatility risk using information from swaptions like Fornari (2008), and creating models for the inflation risk premium like Grishchenko and Huang (2012) would likely be important steps toward improvement.

[33] Preliminary results in a regression of PC1 indicate that the variable is highly statistically significant, has an estimated $\beta \approx -0.23$ and increase R^2 to 67%, but *MFI* loses its statistical significance.

References

Adrian, T., and Wu, H. 2010. "*The term structure of inflation expectations*". Federal Reserve Bank of New York Staff Reports, no. 362. February 2009; revised August 2010.

Ang, A., Bekaert, G., and Wei, M. 2006. *"Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?".* Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Belsley, D. A. 1991. "*Conditional Diagnostics: Collinearity and Weak Data in Regression*". New York: Wiley.

Bernanke, B. 2004. "*What policymakers can learn from asset prices*". Speech before the Investment Analyst Society of Chicago, April 15. (www.federalreserve.gov/ Board-Docs/Speeches/2004/20040415/default.htm).

Campbell, J. Y. and Shiller, R. J. 1991. "*Yield Spreads and Interest Rate Movements: A Bird's Eye View*". Review of Economic Studies 58 (3), 495-514.

Campbell, J. Y., and Shiller, R. J. 1996. "*A Scorecard for Indexed Government Debt*". NBER Working Paper Series, Working Paper 5587.

Campbell, J. Y., and Viceira, L. M. 2001. "*Who should buy long-term bonds?*". American Economic Review, 91(1), 99–127.

Chatterjee, S., and Hadi, A.S. 2012. "*Regression Analysis by Example*". 5th ed. New York: Hoboken, NJ

Chernov, M., and Mueller, P. 2011. "*The term structure of inflation expectations*". Journal of Financial Economics, forthcoming.

Choudhry, M. 2005. "*Understanding the Z-Spread".* Yieldcurve.com, Learning Curve, September 2005.

Choudhry, M. 2006. "*An Introduction To Bond Markets".* Third Edition. John Wiley & Sons, LTD.

Cochrane, J. H., and Piazzesi, M. 2005, "*Bond Risk Premia".* The American Economic Review, Volume 95, No. 1 (March 2005).

D'Amico, S., Kim, D. H., and Wei, M. 2010. "*Tips from TIPS: the informational content of Treasury Inflation-Protected Security prices".* Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Deacon, D., Derry, A., and Mirfendereski, D. 2004. "*Inflation-indexed Securities".* 2nd Edition. John Wiley & Sons, Ltd.

Della Corte, P., Sarno, L., and Thornton, D.L. 2007. "*The Expectations Hypothesis of the Term Structure of Very Short-Term Rates: Statistical Tests and Economic Value*". Journal of Financial Economics.

Ejsing, J., Garcia, J. A., and Werner, T. 2007. "*The Term Structure of Euro Area Break-Even Inflation Rates. The Impact Of Seasonality*". Working Paper Series NO 830, November 2007.

Elsasser, R., and Sack, B. 2004. "*Treasury Inflation-Indexed Debt: A Review of the U.S. Experience*". Federal Reserve Bank of New York Economic Policy Review, May 2004.

Evans, M. D. D. 1998. "*Real rates, expected inflation, and inflation risk premia*". Journal of Finance, 53(1), 187–218.

Fabozzi, F. 2007. "*Fixed Income Analysis 2nd Edition"*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Fisher, I. 1930. "*The theory of interest*". New York: Macmillan.

Fleming, M. J., and Krishnan, N. 2012. "*The Microstructure of the TIPS Market*". Federal Reserve Bank of New York Economic Policy Review, March 2012.

Fornari, F. 2008. "*Assessing the compensation for volatility risk implicit in interest rate derivatives"*. European Central Bank. Working paper series No 859/January 2008.

Grishchenko, O. V., and Huang, J-z. 2012. "*The Inflation Risk Premium: Evidence from the TIPS Market*".

Groβ, J. 2003. "*Linear Regression*". Springer-Verlag Berlin Heidelberg 2003.

Gürkaynak, R. S., Sack, B., and Wright, J. H. 2006. "*The U.S. Treasury Yield Curve: 1961 to the Present"*. Finance and Economics Discussion Series, Division of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Gürkaynak, R. S., Sack, B., and Wright, J. H. 2008. "*The TIPS Yield Curve and Inflation Compensation"*. Finance and Economics Discussion Series, Division of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Hördahl, P. 2008. "*The inflation risk premium in the term structure of interest rates*". BIS Quarterly Review, September 2008.

Ilmanen, A. 1995. "*Time-Varying Expected Returns in International Bond Markets"*. Journal of Finance, Volume 50, No. 2 (Jun., 1995), pp. 481-506.

Ilmanen, A. 2011. "*Expected Returns"*. John Wiley & Sons Ltd.

Karpoff, J. M. 1987. "*The Relation Between Price Changes and Trading Volume": A Survey*. The Journal of Financial and Quantitative Analysis, Vol. 22, No. 1 (Mar., 1987), pp. 109-126.

Keele, L., and Kelly, N. J. 2005. "*Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables*". Political Analysis 2006 14(2):186-205.

Nichols, D. M., and Gonczy, A. M. L. 1994. "*Modern Monetary Mechanics*". Public Information Center, Federal Reserve Bank of Chicago.

O'Brien, R. M. 2007. "*A Caution Regarding Rules of Thumb for Variance Inflation Factors*". Quality & Quantity (2007) 41:673–690.

O'Kane, D. 2000. "*Introduction to Asset Swaps"*. Lehman Brothers International, European Fixed Income Research, Analytical Research Series.

Pflueger, C. E., and Viceira, L. M. 2011. "*Inflation-Indexed Bonds and the Expectations Hypothesis*". The Annual Review of Financial Economics. 3:139-158.

Pflueger, C. E., and Viceira, L. M. Working paper, 2013. "*Return Predictability in the Treasury Market: Real Rates, Inflation, and Liquidity*".

Rencher, A. C. 2002. "*Methods of Multivariate Analysis".* 2nd Edition John Wiley & Sons, Inc.

Rezende, R. B. 2013. "*Risks in macroeconomic fundamentals and bond return predictability*".

Sharma, S. 1996. "*Applied Multivariate Techniques".* John Wiley & Sons, Inc.

Stock, J. H., and Watson, M. W. 1999. *"Forecasting Inflation".* NBER Working Paper No. 7023.

Stock, J. H., and Watson, M. W. 2002. "*Forecasting Using Principal Components From a Large Number of Predictors*". Journal of the American Statistical Association December 2002, Vol. 97, No. 460, Theory and Methods.

Stock, J. H., and Watson, M. W. 2003. "*Forecasting Output and Inflation: The Role of Asset Prices*". Journal of Economic Literature Vol. XLI (September 2003) pp. 788-829.

Trolle, A. B., Schwartz, E. S. 2010. "*An Empirical Analysis of the Swaption Cube*". NBER Working Paper No. 16549.

Tsay, R. S. 2010. "*Analysis of Financial Time Series".* John Wiley & Sons, Inc., Hoboken, New Jersey.

Weill, P.-O. 2007. "*Liquidity Premia in Dynamic Bargaining Markets*". Journal of Economic Theory, Volume 140, Issue 1, May 2008, Pages 66–96.

Wooldridge, J. M. 2013. "*Introductory Econometrics A Modern Approach*". 5th Edition.

Electronic Resources:

Bank for International Settlements, Derivatives Statistics. Available online at: http://www.bis.org/statistics/derstats.htm.

Bureau of Labor Statistics. 2007. BLS Handbook of Methods, Chapter 17, The Consumer Price Index. Available online at: http://www.bls.gov/opub/hom/pdf/homch17.pdf

Federal Reserve Bank of New York, Treasury Auctions. Available online at: http://www.newyorkfed.org/aboutthefed/fedpoint/fed41.html.

Treasury Borrowing Advisory Committee of the Bond Market Association, recommendation to the Secretary of the Treasury May 2001. Available online at: http://www.treasury.gov/resource-center/data-chart-center/quarterly-refunding/Documents/rpt-2001-q2.pdf (2014-04-30).

Appendices

Appendix A

Figure 2: Liquidity Variables

In order to provide a better overview of the liquidity variables, they have not been standardized in the plots below. The off-the-run spread, i.e. the difference in yields between off-the-run and on-the-run nominal bonds, was at its highest during the recent credit crisis. Due to its short history, the relative Z-spread (in basis points) is constant before June 2003 (set to equate to first monthly value of the series). The relative transaction volume (a log-measure) is a moving average over the past 3 months for the ratio in transaction volume for inflation-indexed and nominal debt of similar maturities (approximately 10 year). Lastly, the turnover ratio is a measure for changing liquidity conditions in TIPS, constructed as the accumulated transaction volume within a month for a TIPS-series divided by the total amount of outstanding debt linked to inflation at the end of that month.

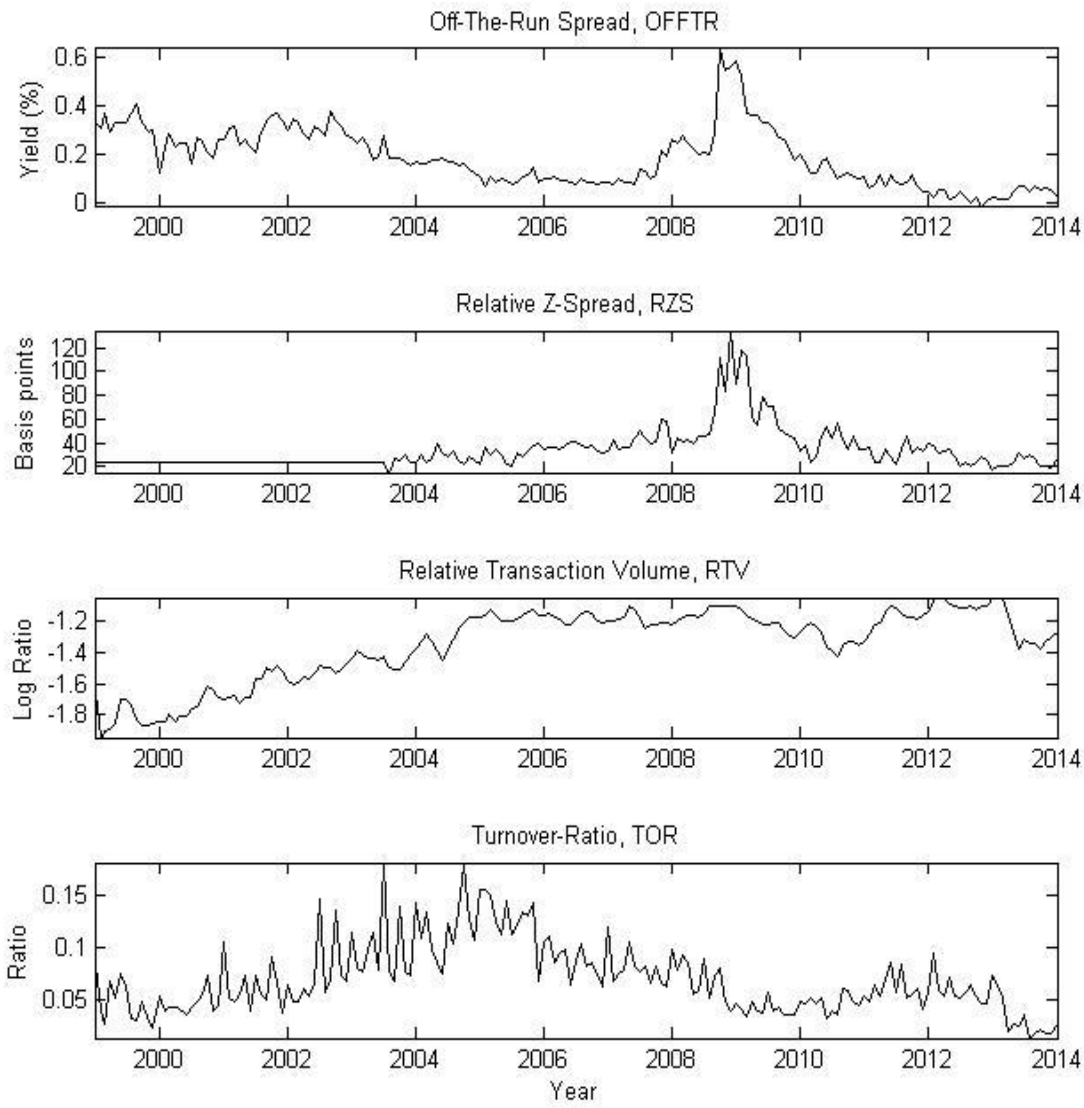


Figure 3: Total amount of outstanding marketable debt

The y-axes indicate the total outstanding amount of debt in billions of dollars. Inflation-indexed debt is shown on the left y-axis, whereas nominal notes and bonds are displayed on the right y-axis. T-bills have been excluded in order to make the numbers comparable.

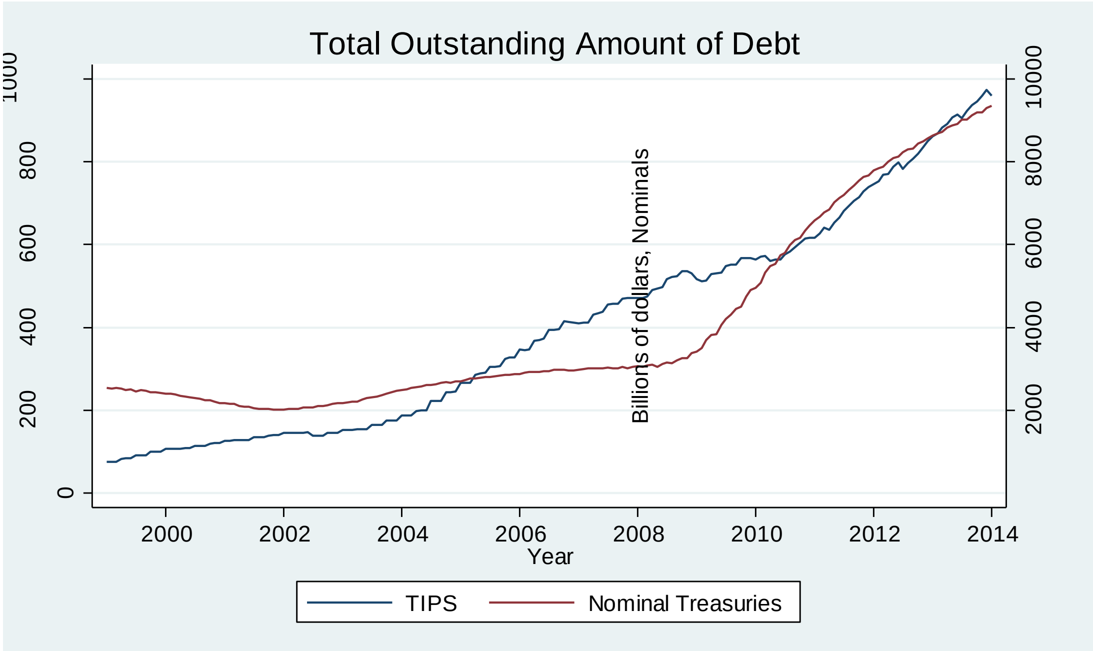


Figure 4: Surveys of forecasted inflation

This figure displays the short- and long-term forecasts for inflation from the Survey of Professional Forecasters and the University of Michigan.

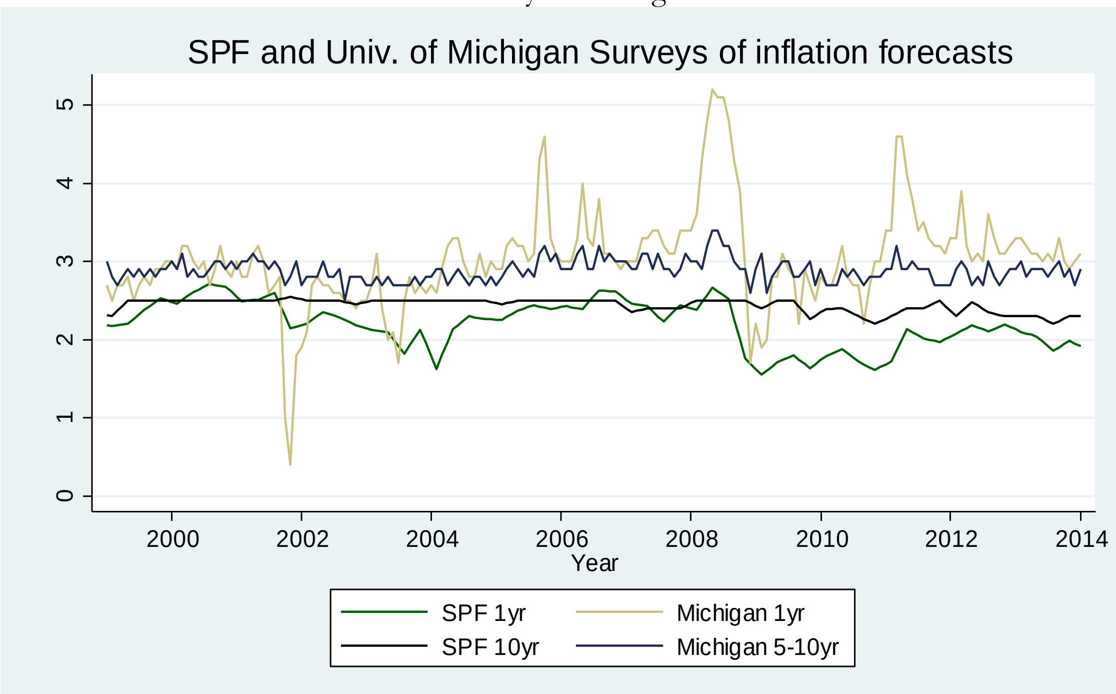


Figure 5: Volatility and inflation risk variables

The figure shows the MOVE index, an index of the implied volatilities of options on nominal US Treasury bonds, and the monthly volatility of the CPI-U computed over a rolling 12-month period. Both of the variables are plotted over the time period from 1999 to 2014.

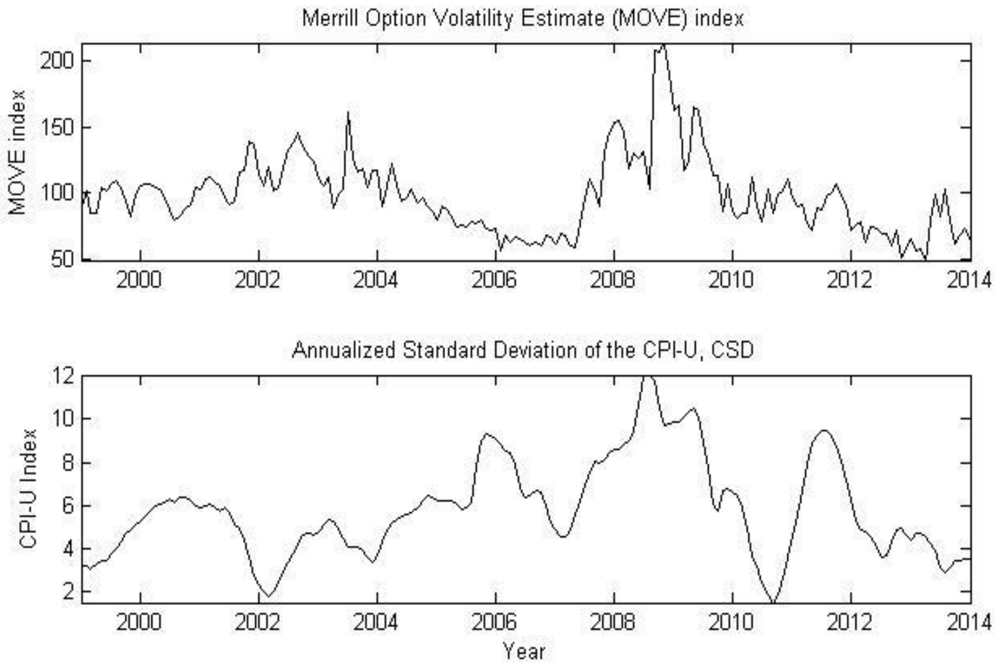


Figure 6: Time series of the first three principal components

This figure displays the first three principal components. In order to more fully appreciate their variation over time, the principal components presented here have not been standardized (doing so makes the series much more alike in a graphical representation).

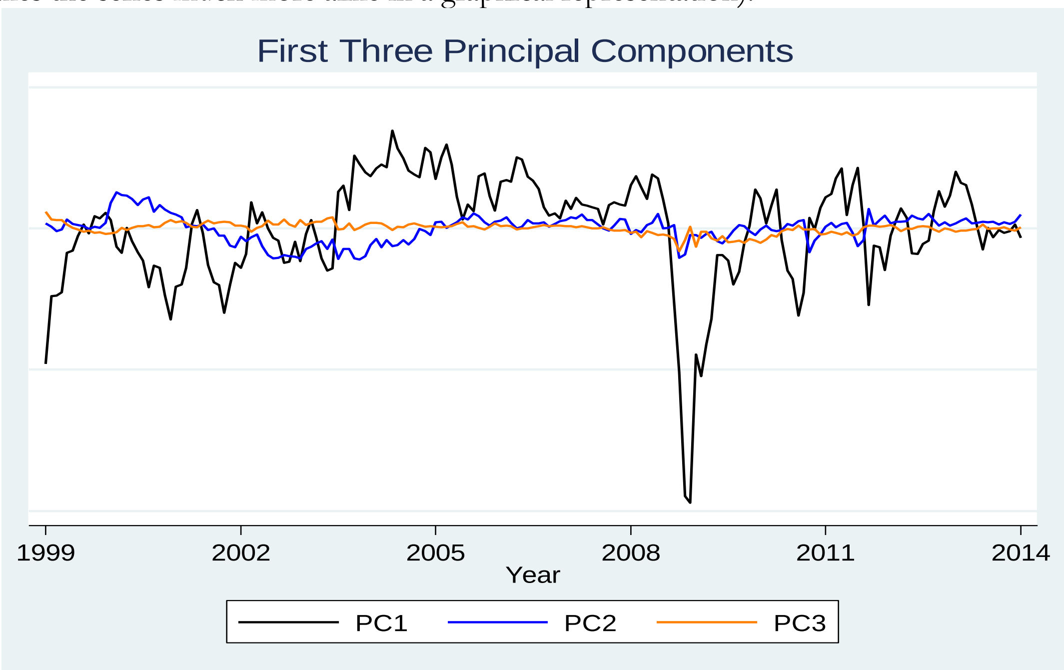


Figure 7: Time series for 5-, 10- and 20-year bond yields and 10-year BEI

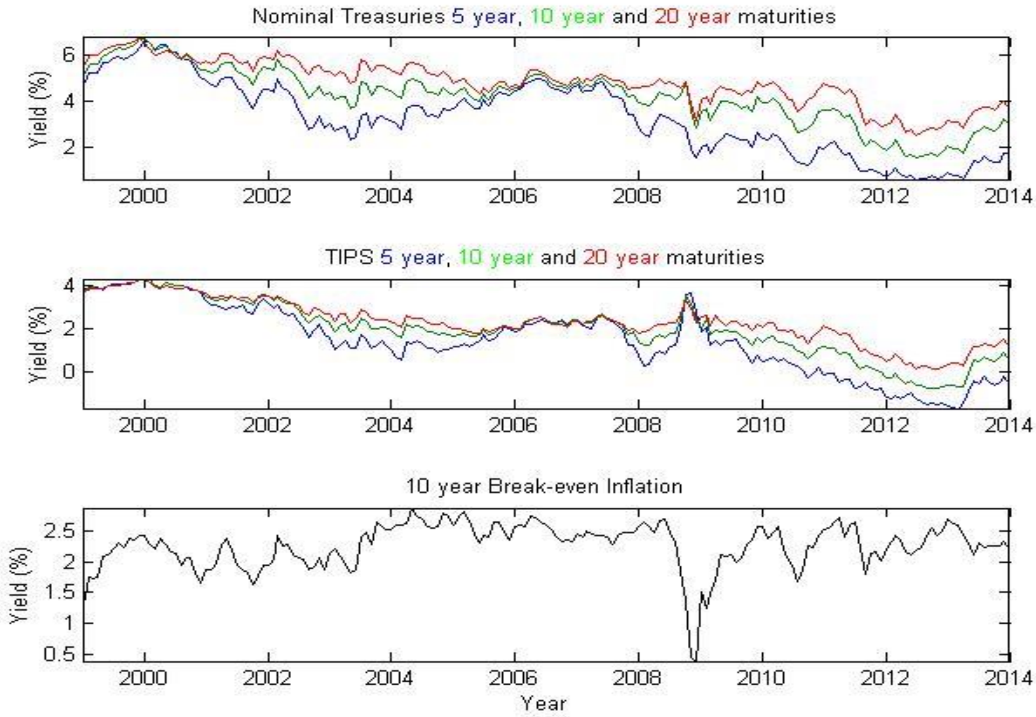


Figure 8: Residuals

This figure displays residuals obtained from running Equation (1).

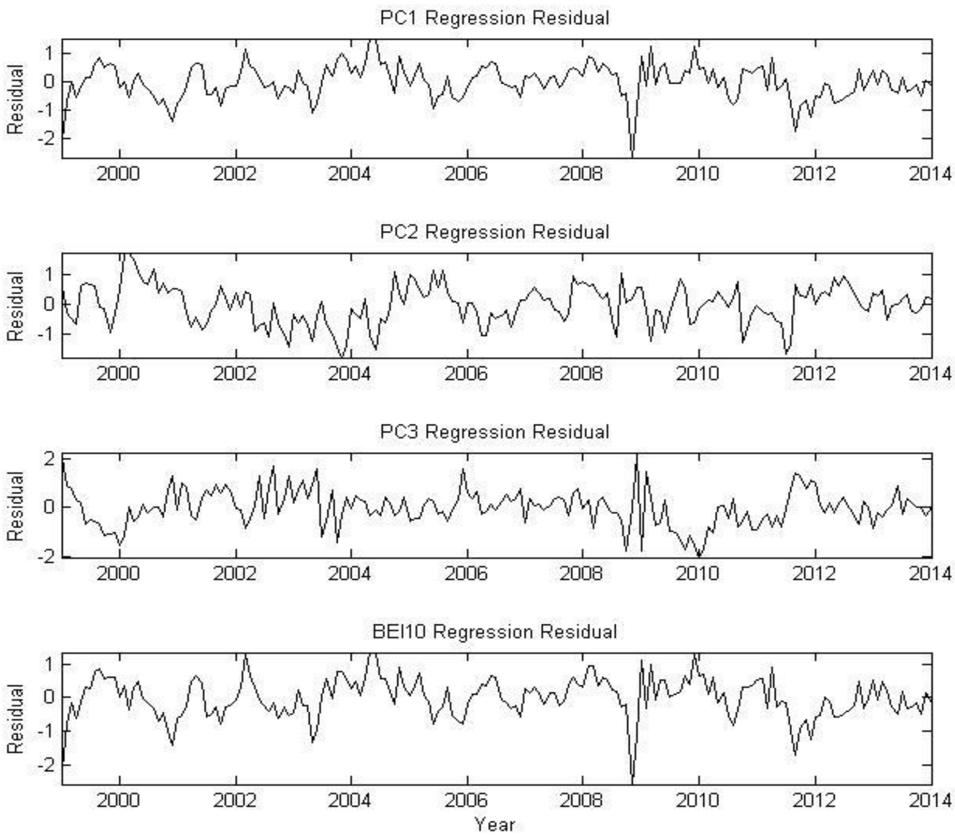


Figure 9: Yield Curves

The figure displays the yield curves of TIPS and break-even inflation respectively at 31 December 2013.[34] The x-axes show the different maturities, while the y-axes show the yields in percent.

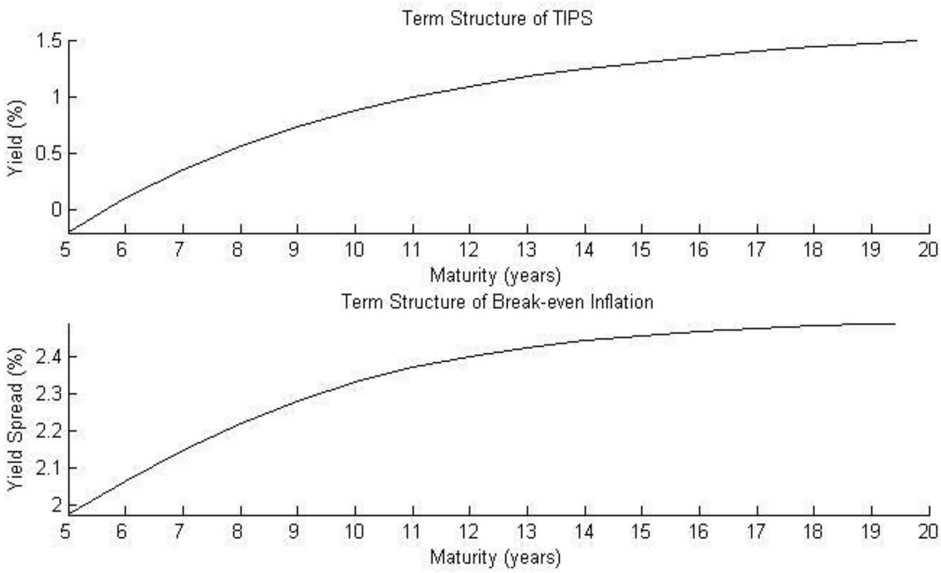
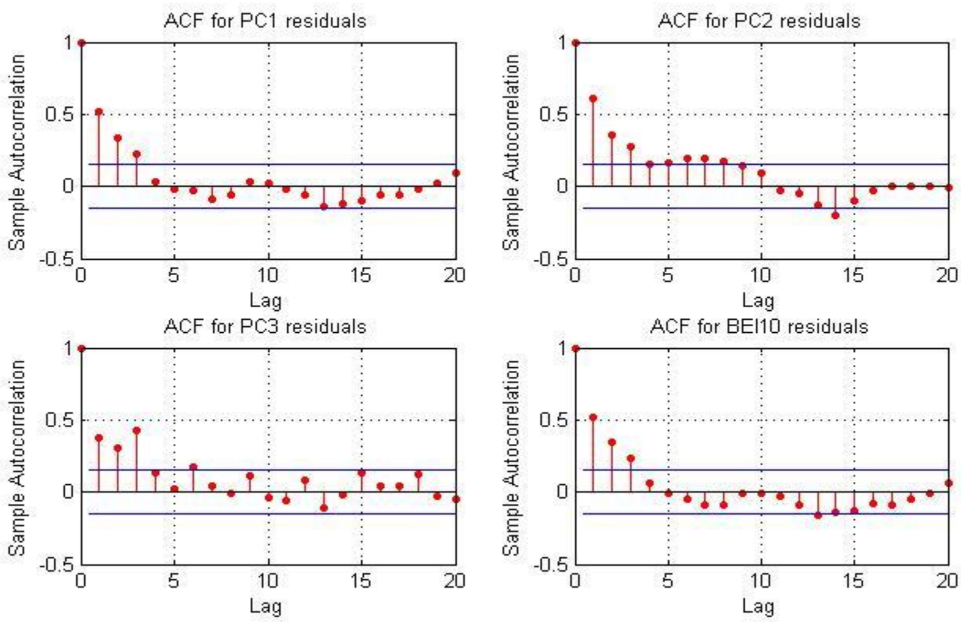


Figure 10: Autocorrelation function (ACF) for residuals

The sample autocorrelation is a statistic that estimates the theoretical autocorrelation. For a series $y_1, y_2, \ldots, y_T$ the sample lag-l autocorrelation is given by: $\hat{p}_l = \frac{\sum_{t=l+1}^{T}(y_t-\bar{y})(y_{t-l}-\bar{y})}{\sum_{t=1}^{T}(y_t-\bar{y})^2}$. Approximate 95% confidence bounds are drawn.



[34] The graphs are based on data from Gürkaynak, Sack and Wright (2006, 2008).

Table 8

Breusch-Pagan/Cook-Weisberg tests for heteroskedasticity

Break-even inflation is abbreviated "BEI". Tests follow the regression of equation (1), where dependent and independent variables alike have been standardized to a mean of 0 and a standard deviation of 1. However, in order to test for the presence of heteroskedasticity, the tests cannot, of course, follow a regression which is specified with robust standard errors. The null hypothesis in each of the tests is a constant variance (homoskedasticity). Test 1: Original Breusch-Pagan/Cook Weisberg test. Assumes that the regression disturbances are normally distributed. Test 2: Computes the N*R2 version of the score test that drops the normality assumption. Test 3: Tests the explanatory variables of the fitted regression model. Test 4: Computes the F-statistic version that drops the normality assumption.

	PC1				PC2			
	Test 1	Test 2	Test 3	Test 4	Test 1	Test 2	Test 3	Test 4
chi2/F	19.34	11.54	28.16	12.19	0.79	0.65	10.6	0.64
Prob>chi2/F	0.00	0.00	0.00	0.00	0.37	0.42	0.23	0.42
	PC3				BEI10			
	Test 1	Test 2	Test 3	Test 4	Test 1	Test 2	Test 3	Test 4
chi2/F	9.28	6.80	51.86	6.98	21.64	12.85	29.95	13.68
Prob>chi2/F	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00

Table 9

Correlation of explanatory variables

Correlation between the regressors used in Equation (1). Monthly data, January 1999 to January 2014 (181 observations).

	MFI	**CFNAI**	**OFFTR**	**RZS**	**RTV**	**TOR**	**MOVE**	**CSD**
MFI	1.00							
CFNAI	-0.10	1.00						
OFFTR	-0.07	-0.53	1.00					
RZS	0.08	-0.64	0.42	1.00				
RTV	0.14	-0.16	-0.45	0.39	1.00			
TOR	0.01	0.18	-0.12	-0.15	0.25	1.00		
MOVE	-0.06	-0.59	0.77	0.52	-0.13	-0.03	1.00	
CSD	0.35	-0.30	0.07	0.53	0.49	0.26	0.09	1.00

Table 10

Regression Results Individual Series

Break-even inflation is abbreviated "BEI". Regressions follow the specification $(PCY_t) = \beta * BEIXyr_t + \epsilon_t$, where Y in PCY takes the values 1,2 and 3, and X in BEIXyr takes on the values 5,10, 15 and 20. This develops into twelve different regressions. *t* indicates the month, going from January 1999 to January 2014. Robust standard errors are used and no constant is used - having standardized the variables to have a mean of 0 (and a standard deviation of 1) makes any intercept redundant. The asterisks (*) after each coefficient signal the significance level of the result. *, ** and *** represent the 1, 5 and 10% significance level respectively. Robust standard errors are stated in brackets below the coefficients.

		PC1		
Factor				
BEI5yr	0.90*			
	(0.04)			
BEI10yr		0.99*		
		(0.01)		
BEI15yr			0.98*	
			(0.02)	
BEI20yr				0.95*
				(0.03)
R^2	0.81	0.97	0.96	0.90
		PC2		
Factor				
BEI5yr	0.40*			
	(0.06)			
BEI10yr		0.13***		
		(0.07)		
BEI15yr			-0.19***	
			(0.10)	
BEI20yr				-0.26**
				(0.12)
R^2	0.16	0.02	0.04	0.07
		PC3		
Factor				
BEI5yr	0.17			
	(0.11)			
BEI10yr		-0.11		
		(0.12)		
BEI15yr			-0.04	
			(0.10)	
BEI20yr				0.16
				(0.10)
R^2	0.03	0.01	0.00	0.03

Table 11

Regression Results Dummies for Crisis

Break-even inflation is abbreviated "BEI". Regressions follow Equation (1) and the period is from January 1999 to January 2014. Dependent and independent variables alike have been standardized to a mean of 0 and a standard deviation of 1. No constant is used - having standardized the variables makes any intercept redundant. The asterisks (*) after each coefficient signal the significance level of the result. *, ** and *** represent the 1, 5 and 10% significance levels respectively. In order to facilitate the reader in discerning the effects of the dummies, robust standard errors are (in contrast to other tables showing regression results) not stated in brackets below the coefficients. The two variables *MOVE* and *CSD* are excluded for brevity because here we only desire to investigate whether the crisis around 2008 impacted liquidity to a significant extent in our sample. For each dependent variable, the first column contains the results from regressing only onto the four liquidity proxies. Thereafter the second column and third column for each regressand display the outcome of also including the dummies ("_dum") for the 2008 crisis and the measures for short- and long-term expected inflation respectively. The estimates of the dummies are in bold.

	PC1			**PC2**		
Factor	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
MFI			0.10**			0.28*
CFNAI			0.36*			-0.16**
OFFTR	-0.38*	-0.30*	-0.14	-0.82*	-0.98*	-0.98*
RZS	-0.36*	-0.16	-0.19**	0.43*	0.52*	0.41*
RTV	0.11	0.11	0.22**	-0.55*	-0.71*	-0.70*
TOR	0.28*	0.27*	0.22*	-0.15**	-0.12	-0.11
OFFTR_dum		**-0.57**	**-0.62**		**0.29**	**0.36**
RZS_dum		**-0.20**	**-0.08**		**-0.43****	**-0.18**
RTV_dum		**0.99***	**1.25***		**1.18***	**0.10**
TOR_dum		**0.11**	**-0.08**		**0.01**	**0.20**
R^2	0.55	0.59	0.65	0.32	0.36	0.43

	PC3			**BEI10yr**		
Factor	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
MFI			0.00			0.13**
CFNAI			-0.18***			0.35*
OFFTR	-0.29**	-0.11	-0.18	-0.45*	-0.41*	-0.24**
RZS	-0.13	-0.44*	-0.44*	-0.28**	-0.04	-0.08
RTV	-0.50*	-0.29***	-0.33**	0.10	0.06	0.16***
TOR	0.34*	0.31*	0.33*	0.21*	0.21*	0.16*
OFFTR_dum		**-0.92***	**-0.89***		**-0.42**	**-0.47**
RZS_dum		**1.03***	**1.02***		**-0.36**	**-0.21**
RTV_dum		**-1.01***	**-1.31***		**1.19***	**1.35***
TOR_dum		**-0.22**	**-0.11**		**0.16**	**-0.02**
R^2	0.36	0.45	0.46	0.51	0.57	0.63

Table 12

Uncentered Variance Inflation Factors (VIFs)

This is a postestimation test conducted after estimates of Equation (1) has been obtained. The dependent variable does not matter since it is the independent variables that the variance inflation factors are being calculated for. 1/VIF is also known as tolerance level.

Variable	VIF	1/VIF
OFFTR	5.35	0.19
RTV	3.69	0.27
RZS	3.22	0.31
MOVE	3.09	0.32
CFNAI	2.32	0.43
CSD	2.13	0.47
TOR	1.38	0.72
MFI	1.27	0.79
Mean VIF	2.81	

Table 13

Tests for Serial Correlation

Break-even inflation is abbreviated "BEI". Standardized series. Tests follow the regression of equation (1), however not specified with the option for robust standard errors since this would interfere with the tests. Test 1: Durbin-Watson d statistic to test for first-order serial correlation in the disturbance when all the regressors are strictly exogenous. Test 2: Durbin's alternative test for serial correlation in the disturbance. This test does not require that all the regressors be strictly exogenous. Test 3: Breusch-Godfrey test for higher-order serial correlation in the disturbance. This test does not require that all the regressors be strictly exogenous. The null hypothesis in Dubin's alternative test and the Breusch-Godfrey test for higher-order serial correlation is no serial correlation. For a sample size of 200 and 8 explanatory variables (with no intercept), the Durbin-Watson D-statistic lower limit (dL) is 1.582 and the upper limit (dU) is 2.233 serial correlation (on a 1% level). Values falling below the lower limit are evidence of serial correlation.

	PC1			PC2		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
D-statistic (8, 181)	0.89			0.78		
chi2		73.45	54.17		108.86	70.16
Prob>chi2		0.00	0.00		0.00	0.00

	PC3			BEI10		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
D-statistic (8, 181)	1.19			0.88		
chi2		36.79	31.90		75.71	55.32
Prob>chi2		0.00	0.00		0.00	0.00

Table 14

Tests for Serial Correlation with lags on dependent variable included

Break-even inflation is abbreviated "BEI". Standardized series. Tests follow the regression of equation (1), however not specified with the option for robust standard errors since this would interfere with the tests and we have also included lags (either 1 or 2) for the dependent variables as explanatory variables. Test 1: Durbin's alternative test for serial correlation in the disturbance. Test 2: Breusch-Godfrey test for higher-order serial correlation in the disturbance. Neither of the two tests require that all the regressors be strictly exogenous. The null hypothesis in Dubin's alternative test and the Breusch-Godfrey test for higher-order serial correlation is no serial correlation. Including lags of the dependent variable rules out the use of the Durbin-Watson d statistic.

	PC1				PC2			
	1 lag		2 lags		1 lag		2 lags	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
chi2	19.12	18.20	2.12	2.23	1.07	1.12		
Prob>chi2	0.00	0.00	0.15	0.14	0.30	0.29		

	PC3				BEI10			
	1 lag		2 lags		1 lag		2 lags	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
chi2	4.66	4.80	0.53	0.57	18.22	17.43	3.69	3.85
Prob>chi2	0.03	0.03	0.47	0.45	0.00	0.00	0.05	0.05

Table 15

Order of Integration

Results from the Augmented Dickey-Fuller test. The variables (both dependent and independent as used in Equation (1)) have here been standardized to a mean of 0 and a standard deviation of 1.

Variable	I(O)	Variable	I(O)
MFI	0	PC1	0
CFNAI	0	PC2	0
OFFTR	1	PC3	0
RZS	0		
RTV	1		
TOR	1		
MOVE	1		
CSD	0		

Appendix B

Historical summary of indexed securities

As described in Deacon, Derry and Mirfendereski (2004), one of the first known instances of the use of indexed financial instruments dates back to the middle of the 18th century when the State of Massachusetts issued bills of public credit linked to the cost of silver on the London Exchange. However, this first, somewhat experimental, issue had an unfavorable outcome for the State. The price of silver appreciated more than prices of goods in general, leading to a significant increase in the real value of the debt, which subsequently lead to heavy losses for the issuer. The State's next issuance of indexed debt, in the form of *Depreciation Notes*, came in 1780 during the American Revolution. Adhering to a law passed by the State Parliament, this issue would be indexed to a basket of common consumer goods like corn, beef and sheep wool instead of silver. The debt was used as payment to soldiers in lieu of wages.

There has been quite a lot of discussion in the literature about the potential advantages and drawbacks with issuing and buying indexed debt (see for example Campbell and Shiller (1996) and Bernanke (2004)). Throughout the last decades, many sovereign and corporate issuers have sold these instruments. One way issuers may gain from selling inflation-indexed bonds rather than regular (nominal) bonds is by avoiding the cost of having to compensate buyers of nominal bonds with an inflation risk premium. Furthermore, if realized inflation turns out to be lower than the market expects, indexed debt becomes the ex post cheaper source of financing compared to nominal bonds.

Another common argument for linking debt to a price index is that governments would have more incentive to keep inflation in check. It would be more costly to conduct inherently inflationary policies like running budget and current account deficits if part of the debt was linked to inflation. Others argue that wide-spread indexation would, on the contrary, signal to the market that politicians have given up and that there is no political will to keep inflation under control. An argument that is often discussed in the literature, and which was scrutinized in the 1990s before the US Treasury first introduced its inflation-indexed debt, is that these instruments are valuable as tools to policy-makers like the Federal Reserve. Until the introductions of these instruments there was no way to observe real rates in the market, and therefore, market participants' expectations about future inflation could not be derived either.

Having an asset that provides a stable real rate of return with no credit risk would also be good, many argue, since people seeking to invest their savings would be able to earn a decent return which keeps up with the increases in the cost of living. Historically, one can observe that growth in developed economies have correlated well with inflation, and vice versa. Inflation-indexed debt would help smooth government deficits (and surpluses) out over time as strong growth generally would result in more tax revenues, making it easier to service the higher nominal debt that would be incurred because of higher inflation. In bad times, nominal debt payments would go down, thus relieving the economy a bit from the interest on the debt.

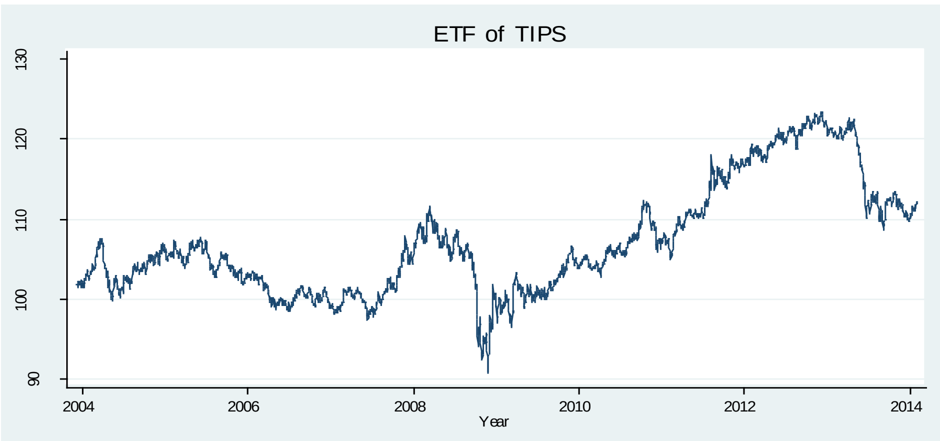
Over the years, many influential economists like Alfred Marshall, Irving Fisher, John Maynard Keynes and Milton Friedman have argued for a wider adoption of indexed debt. Indexed bonds started becoming popular only after the period of high and volatile inflation in many economies during the second half of the 20th century. France and Finland issued indexed debt as part of their price stabilization programs following the Second World War, and during the 1950s and 1960s, several South American countries experiencing hyperinflation issued these securities as a way of resuscitating a market for long-term debt. More recent issuers have largely seen indexed debt as a source of cost saving and as a way to further increase the credibility of their monetary policy. Many of the world's major economies just started issuing indexed debt in the last 30 years, e.g. the UK, Australia, Canada, Sweden, USA, France, Greece and Italy.

Historical real asset returns

How does the historical risk and returns of real assets measure up to nominal assets? Data on real bond returns is quite scarce. However, there exists an ETF that mimics the performance of a portfolio TIPS. The historical performance for the iShares TIPS bond ETF, which directly owns TIPS bonds, is shown below. While this measure of TIPS returns is not perfect, it should give us a good approximation. The majority of the bonds that this ETF owns have maturities between 1 to 10 years.

Figure 11: iShares TIPS bond ETF

This figure presents the price history for the iShares TIPS bond ETF for the period 2004-2014.[35] The y-axis shows the price of the ETF, and the x-axis the date.



[35] The data is obtained from Yahoo! Finance.

Additionally, in Table 16 we present average historical returns on some nominal assets that are then compared to the average change in the CPI. Finally, in Table 17, we present some historical real changes in asset prices and macroeconomic variables.

Table 16[36]

Historical risk and return on stocks and nominal bonds

Asset returns, volatility and correlations to equity returns for the period 1990-2009.

Asset	**Geometric mean (%)**	**Annual volatility (%)**	**Correlation to equity**
Global developed equity markets	5.9	15.4	1
US Treasuries	6.8	4.7	-0.03
Treasury Bill	3.8	0.6	0
CPI	2.8	0.9	-0.07

From the table above we can compute that the average real Treasury return over the period was 3.9% and the average real Bill return was 1.0%.

Table 17[37]

Historical real returns on assets and price changes of macroeconomic variables

Average annual real returns on assets and real price changes in various macroeconomic indicators for the 1900s in the US.

Asset	**Real return (%)**
Stock market	7.5
Treasury bills	0.9
Treasury bonds	1.5

Macroeconomic indicator	**Real price change (%)**
House prices	0.3
Oil price	-0.3
Real short rate	0.9
Ex ante 10 year real yield	2.8
Yield curve	0.9
Real GDP growth	3.4
CPI inflation	3.1
Ex ante 10 year inflation	2.2

[36] The table is constructed from data on historical asset returns in Ilmanen (2011).
[37] The table is constructed from data on real asset returns and price changes in Ilmanen (2011).

The real T-bill and Treasury bond returned on average 0.9% and 1.5% per annum respectively during the 20th century. This can be compared to the 1.0% and 3.9% annual real during the two decades ending in 2009. From a historical perspective, this period has seen great real returns on long-term bonds. We also note that long-term rates and inflation have consistently come down and remained at low levels for the last 30 years.

Treasury auctions and trading

Bond auctions, as outlined by the New York Fed[38], are usually held in a single-price auction format. At single price auctions, bidders receive their allocation at the clearing price, but at multiple price auctions, bonds are filled from the highest price down until all bonds have been allocated. Multiple-price auctions usually favor more sophisticated investors whom likely have a better understanding of how to price new instruments. As a consequence, that type of auction may discourage the participation of smaller investors.

Treasury securities are first sold in the primary market through single-price auctions held on a set schedule. The process begins with a public announcement of the auction details by the Treasury Department a couple of days in advance. The two types of bids that are accepted are non-competitive tenders and competitive bids. Non-competitive tenders are guaranteed to receive securities, but the maximum amount that may be sold to a single non-competitive bidder is small. In general, most Treasuries are bought by primary dealers, and the volume bought through non-competitive tenders is usually very small. When allocating the securities among bidders, the Treasury first grants securities to the non-competitive tenders and then, starting with the highest bid, work their way through the list of competitive bids until all the securities are allotted. The bonds are then all sold at the highest accepted yield, and a fixed coupon rate determined through the auction process is set on each new issue. Secondary market trading takes place in the over-the-counter market where dealers continuously quote bid and ask prices on outstanding Treasury issues. When-Issued trading (trading in bonds that have been announced but not yet issued) helps price discovery and increases transparency for those that will bid at auctions. Bonds are often issued over a period of time. This method of issue is called tap issue. For instance, the US Treasury provides a tentative auction schedule of future Treasury security auctions. As an example of how trading volume in TIPS has increased over the years, the weekly trading volume went from $900MM in 1998 to $12,000MM in 2013.

[38] See references for further information.

TIPS pricing

As given by Deacon, Derry and Mirfendereski (2004) the price P_R for an idealized real bond, which is perfectly indexed, paying real coupons c_R at time t and the real redemption payment M_R at maturity T is given by:

$$P_R = \sum_{t=1}^{T} \frac{c_R}{(1+r_t)^t} + \frac{M_R}{(1+r_T)^T}$$

for the formula to hold exactly though, it is also necessary to assume that the first coupon payment will be paid in exactly one period's time to avoid making an adjustment for accrued interest.

Attentive readers might notice that the formula above does not differ from the pricing formula of a regular nominal bond, and that this is due to the perfect indexation assumption, i.e. that the measure of inflation used to scale up the real cash flows is the same as the one that is used to scale up the discount rate. Unfortunately, in reality, it is not this simple. To compute a real yield for bonds that have an indexation lag that is equal to the coupon period we have to make an assumption about future inflation rates:

$$P_R = \sum_{t=1}^{T} \frac{c_R \prod_{i=0}^{t-1}(1+\pi_i)}{(1+r_t)^t \prod_{i=1}^{t}(1+\pi_i)} + \frac{M_R \prod_{i=0}^{T-1}(1+\pi_i)}{(1+r_T)^T \prod_{i=1}^{T}(1+\pi_i)}$$

which can be simplified to

$$P_R = \sum_{t=1}^{T} \frac{c_R(1+\pi_0)}{(1+r_t)^t(1+\pi_t)} + \frac{M_R(1+\pi_0)}{(1+r_T)^T(1+\pi_T)}$$

where π_t is the realized inflation rate between time t-1 and t and r is the real discount rate. Now, let us set $r_1 = r_2 = \cdots = r_T = \bar{r}$ and $\pi_1 = \pi_2 = \cdots = \pi_T = \bar{\pi} = 2\%$ (for example). Given π_0, P_R, c_R and M_R, we can solve the above equation with respect to the real yield-to-maturity $\bar{r}$.

$$P_R = (1+\pi_0) \sum_{t=1}^{T} \frac{c_R}{(1+\bar{r})^t} + \frac{M_R}{(1+\bar{r})^T}$$

The longer the lag and the shorter the residual maturity, the bigger the impact this inflation assumption will have on the computed yield, and the greater an approximation computed yields will represent. Thus, these observed real yields are still only approximations to true real rates. In practice, because indexed bonds in most major markets are issued on a real yield basis, authorities have to publish an official price-yield equation for settlement purposes. For simplicity, these official formulae assume that the expected inflation terms cancel entirely.

The settlement price for TIPS is given by the official US Treasury price-yield formula:

$$P_N = Inflation - adjusted\ price + inflation\ accrual$$

$$Inflation - adjusted\ price\ =\ Real\ price\ *\ Price\ index$$

$$Real\ price = \left(\frac{1}{1+\frac{f}{d}\frac{r}{2}}\right)\left(\frac{c}{2}+\frac{c}{2}\sum_{j=1}^{n} w^j + 100w^n\right)$$

$$Inflation - adjusted\ accrued\ interest\ =\ RAI\ *\ Index$$

where $w = 1/(1+r/2)$, *RAI* is the unadjusted or real accrued interest = $c/2 * (d-f)/d$, f is the number of days from settlement date to the next interest payment date, d is the number of days in the regular semi-annual coupon period ending on the next interest payment date, and n the number of full semi-annual coupon periods between the next interest payment date and the maturity date.

This method uses simple interest for discounting cash flows during the current coupon period and is used by the Treasury to compute settlement proceeds at auction. However, traders in the secondary market compute yields from prices by using compound interest for discounting cash flows.

Further bond risks

Interest rate risk

Real interest rate risk is the risk relating to the uncertainty of future changes in the real rate of interest. For bearing this risk, market participants will require some compensation. This compensation can be called the real interest rate risk premium.

The more common concept of modified duration, which is often simply called "interest rate risk", is defined as the percentage change in a bond's price given an incremental change in its yield. From Fabozzi (2007), we know that bond features that affect interest rate risk include the maturity of the bond, the coupon rate of the bond and the interest rate level (i.e. the yield-to-maturity). Interest rate levels affect interest rate risk through the so-called convexity effect. Positive convexity can easily be illustrated by imagining that the price-yield relationship of a bond is given by a convex function. The greater the convexity of the bond, the less effect a change in interest rates would have on the price of the bond.

By interest rate risk we often only mean the risk of parallel shifts in the yield curve, affecting short-term rates and long-term rates by an equal amount. However, the yield curve might change by different amounts at different maturities along the curve. This is called yield curve risk. In this paper we will not make any distinction between these two different types of interest rate risk, but will instead be alluding to both of them together when discussing interest rate risk. Interest rates, or the price of money, are determined by the forces of supply and demand just like any other prices. We will therefore divide this discussion into two parts. First, we will discuss aspects important for the demand for money, and then we will briefly touch upon the subject of money supply.

Fisher (1930) explains that the demand for money, at a fundamental level, is determined by individuals' time preferences. Time preference is the relative valuation of utility now or in the future, and is dependent solely upon different individuals' subjective preferences. The following examples show more clearly how preferences are connected to the demand for funds: Assume that investment opportunities in an economy are plentiful. Investor will then be demanding funds, which is expressed by a larger demand for savings, to be able to invest in projects offering attractive enough risk-adjusted returns. This demand for money will drive interest rates up. On the other hand, if something unforeseen happens like a natural disaster, people might be forced to make the decision to consume now (rebuild their house) rather than to save and invest. Those people would of course value the immediate utility of having a roof over their head more than the return they would get from saving the money, and the resulting effect would, ceteris paribus, be lower interest rates. It is important to remember that we are only concerned with the effect on interest rates of a change in the demand for funds here, and that we ignore the fact that savings would also have to be spent (which would put upward pressure on interest rates) in order to pay for the reconstructions of the homes.

Time preference risk is the risk that market prices (through discount rates) respond adversely to changes in preferences for savings and investment. Hence, it is also intimately connected to real interest rate risk and hedging demand for real interest rate products. To exemplify this concept, assume that a pension fund with liabilities denominated in real terms (such as the British pension system) has a portfolio of inflation-indexed debt securities with a duration of 20 years that matches its liabilities well. If, for some reason, the time preference in the economy gets higher (higher preference for consumption today), the lower demand for funds farther into the future will make the pension fund sell some of its long-term securities in order to better hedge the new, lower duration of its liabilities. This will result in higher long-term real rates all else equal. An additional example connected to regulatory risk is that rates might change in response to changes in investors' legal mandates and regulation. If, for example, institutional investors that can only invest in investment grade bonds could also invest in bonds with lower credit rating, one would expect that rates on those bonds would go lower.

The supply of money is today regulated by central banks. As discussed in Nichols and Gonczy (1994), by manipulating the supply of money, a central bank can control the short-term rate of interest. In the US, the Federal Reserve sets an interest rate called the Fed Funds Rate which is the rate at which banks can borrow and lend each other excess reserves held at the Fed. To implement monetary policy, the Fed makes use of the following tools: open market operations, the discount rate, bank reserve requirements and forward guidance. The tool predominantly applied to control the Fed Funds Rate is through open market operations whereby the Federal Open Market Committee (FOMC) decides to buy or sell Treasury securities in order to adjust the amount of bank reserves available. They buy securities with money they create ex nihilo (colloquially known as "printing money"). In turn, this creates an equal amount of new bank deposits which can then be used to make loans to businesses and private individuals.

Other bond risks

Another risk that bond investors are often exposed to is credit risk. This is the risk that the issuer of the bond is not able to satisfy the terms of its obligation with respect to the timely payment of interest and principal. Since it is generally accepted that US Treasury securities have no default risk even though the US had their credit rating downgraded in August 2011 by S&P, and that US CDS spreads indicate otherwise (they are not 0!), we will make the assumption that there is no credit risk. For the purposes of this paper we will also ignore currency risk, as well as various regulatory risks and political risks.

Index risk is the uncertainty the market feels about the price index which is used to calculate the cash payments to an inflation-linked bond. Some reasons for such a risk to be present could be related to issues with the reliability, integrity or timing of the index. Potential reliability issues could include reservation about how the index is defined and calculated. If an inappropriate price index is used, the indexed debt using that price index might itself be inappropriate as a hedging instrument. While investors seeking to protect themselves from erosion of purchasing power might prefer a price index based on a basket of consumer goods, a government issuer might prefer to use the GDP deflator because it has a higher correlation with its revenues and expenses. Also, if a consumer price index does not adequately reflect the real changes in the cost of living because of exclusion of important goods or untrustworthy measurements of prices, the price index will also be inappropriate. Other sources of uncertainty could come from potential changes in the composition or weightings in the index, or from changes in how the index is calculated.

An investor holding a bond to maturity will receive the yield to maturity of the bond, assuming that the investor can invest the coupons of the bond at an average rate equal to the yield-to-maturity. The risk of not being able to do this is called reinvestment risk. Since future rates are unknown, every coupon-bearing bond has some reinvestment risk. Bonds with higher coupons have more reinvestment risk than bonds with lower coupons. Zero-coupon bonds have no reinvestment risk. Prepayment risk, the risk that the issuer will call (redeem) the bond before maturity, is not present for the types of bonds that we will study in this thesis.

The term premium is the compensation the market requires for being exposed to the extra risks involved in owning long-term debt securities. There are many competing explanations about how the term structure of interest rates behaves, one such prospective explanation is the expectations hypothesis. It can be simplified as stating that today's long-term rates are just expected future short-term rates. There has been a lot of research on the expectations hypothesis and it does not seem to hold up to empirical scrutiny, see Campbell and Shiller (1991) and Della Corte, Sarno and Thornton (2007). The liquidity preference theory can be simplified as asserting that the extra risk comes from the fact that economic uncertainty increases the farther into the future you look, and from that the price of a long-term debt security is more sensitive to changes in interest rates than the price of a short-term debt security (i.e. long-term debt has higher duration than short-term debt). Proponents of the market segmentation theory believe that the yield curve is simply a reflection of the supply and demand for funds at different maturities, and that the different interest rates along the yield curve are determined independently of each other.

Principal Component Analysis

As discussed in Sharma (1996), Tsay (2010) and Rencher (2002), the PCA method is used as a tool for dimensionality reduction of data when there is correlation among observations. The technique is concerned with forming a set of new uncorrelated variables called principal components from the original data. Ideally, a small number of these principal components, fewer than the number of original variables, should be able to capture most of the variation in the data.

Let $\boldsymbol{R} = (\boldsymbol{r}_1, \boldsymbol{r}_2 \ldots, \boldsymbol{r}_n)$ be an n-dimensional random variable with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}}$. The principal components ξ_i are linear combinations of $\boldsymbol{r}_i$ that summarize the structure of $\boldsymbol{\Sigma}_{\boldsymbol{r}}$. The first principal component is constructed to account for maximum variance in the data. The second principal component is constructed so that it accounts for the maximum amount of variance in the data not explained by the first principal component, and so on. Each principal component is orthogonal to every other principal component.

Expressing this algebraically, we want to find a weight matrix w_{ij} for the following system of n equations such that they maximize Var(ξ_i):

$$\boldsymbol{\xi}_i = \boldsymbol{w}_i \boldsymbol{r} = \sum_{j=1}^{n} w_{ij} r_j$$

subject to the constraints

$$\boldsymbol{w}_i \boldsymbol{w}_j' = 1$$

and

$$\forall\ i \neq j\text{: } \mathrm{Cov}\left(\xi_i, \xi_j\right) = 0$$

Another approach is to use spectral decomposition on the sample covariance matrix to find the matrix of eigenvectors representing w_{ij}, and where the corresponding eigenvalues show the fraction of explained variance to total variance in the data that each principal component contributes with.