Decomposing Long Bond Returns: A Decentralized Modeling Approach

Peter Carr

Tandon School of Engineering, New York University

Liuren Wu

Zicklin School of Business, Baruch College, The City University of New York

Abstract

This paper develops a decentralized theory that determines the fair value of the yield-to-maturity of a bond or bond portfolio based purely on the near-term dynamics of the yield itself. The theory decomposes the yield into three components: its expected change, its risk premium, and its convexity effects. The convexity effect can be constructed with a historical variance estimator. The expectation can come from statistical models or expert forecasts, leaving the remaining component of the yield as a risk premium estimate. Comparative yield analysis of different bonds can start with commonality assumptions on their risks, risk premiums, and expected change.

JEL Classification: C13, C51, G12, G13

Keywords: Bond returns, profit and loss attribution, yield decomposition, expectation, risk premium, convexity effects, butterfly trades

The authors thank Zhenyu Cui, Ionut Florescu, Nicola Fusari, Scott Joslin, and seminar participants at the Financial Engineering Seminar at Stevens Institute of Technology, the 2016 JHU-QQR Conference on "The Role of Derivatives in Asset Pricing," the 2017 NYU Conference on "Derivatives and Volatility 2017: The State of the Art," and the 2018 American Finance Association Meetings for their comments and suggestions. Liuren Wu gratefully acknowledges the support by a grant from The City University of New York PSC-CUNY Research Award Program. Address correspondence to: Liuren Wu, Department of Economics and Finance, Zicklin School of Business, Baruch College, CUNY, One Bernard Baruch Way, Box B10-225, New York, NY 10010, (646) 312-3509 or email: liuren.wu@baruch.cuny.edu.

Contents

The literature on interest rate modeling is vast, with different approaches targeting different challenges. The traditional literature on expectation hypothesis (e.g., Campbell and Shiller (1991), Campbell (1995), and Bekaert, Hodrick, and Marshall (1997)) focuses on predicting changes in short-term interest rates with the yield curve slope. The yield curve shape contains information on not only the expectation but also risk premium and convexity effects, but the expectation component dominates the short end of the yield curve (Longstaff (2000)) and can be used to predict future short rate movements. During the past decade, the literature on no-arbitrage dynamic term structure models (DTSMs) has experienced tremendous growth in both theoretical characterization and empirical analysis.[1] By specifying the instantaneous interest rate dynamics and applying the principle of no dynamic arbitrage, these models generate fair values on the whole yield curve and are thus capable of pricing bonds of all maturities within one centralized view of the short rate dynamics. The centralization has played important roles in practical applications, such as generating interpolated valuation in between observed maturities and identifying relative valuation opportunities (e.g., Duarte, Longstaff, and Yu (2007) and Bali, Heidari, and Wu (2009)). Finally, to price interest rate options, the literature often takes the observed yield curve as given and focuses on modeling the interest rate volatility.[2] This treatment highlights the contribution of interest rate volatility to option valuation while delta hedging the directional interest rate exposure.

All these existing frameworks, however, have limited success in explaining the short-term return behavior of long-dated bonds. The expectation hypothesis literature uses long rates to predict

[1] See Duffie and Kan (1996), Duffie, Pan, and Singleton (2000), and Duffie, Filipović, and Schachermayer (2003) for a progressively more general characterization of the affine class of models, Leippold and Wu (2002) for a characterization of the quadratic class, and Cheng and Scaillet (2007) for an integration of the two classes. Prominent examples for empirical analysis include Dai and Singleton (2000), Dai and Singleton (2002), Duffee (2002), Leippold and Wu (2003), Aït-Sahalia and Kimmel (2010), Adrian, Crump, and Moench (2013), Hamilton and Wu (2012), Collin-Dufresne, Goldstein, and Jones (2008), Joslin, Singleton, and Zhu (2011).

[2] Examples include the forward rate models of Ho and Lee (1986) and Heath, Jarrow, and Morton (1992); the market rate models of Brace, Gatarek, and Musiela (1997), Jamshidian (1997), and Miltersen, Sandmann, and Sondermann (1997); the string models of Goldstein (2000) and Santa-Clara and Sornette (2001); and the nonparametric pricing approach of Aït-Sahalia (1996).

short rate changes, not the other way around. Modeling long rates with DTSMs also stretches the modeler's imagination on how short rate should move in the far distant future. The starting point for this literature is often some mean-reverting dynamics assumption on the short rate; yet, mean reversion calibrated to the short end of the yield curve often implies much smaller movements for long rates than actually observed from data (Giglio and Kelly (2018)), and asymptotic claims based on stationarity assumptions (e.g., Dybvig, Ingersoll, and Ross (1996)) often run counter to actual observations on the behavior of very long-dated bond yields. Long rates are neither easily predictable, nor converging to a constant any time soon. They tend to move randomly, with substantial volatility, a behavior that is difficult to be reconciled within most DTSM specifications. Furthermore, the centralized nature of DTSMs can experience error-contagion issues in practical implementation: Adding or removing one security from the estimation or an accidental data error on one security can alter the fair valuations on all other securities.

It is important to realize that investors can choose to hold a very long-dated bond for a very short period of time. In this case, investors are mainly concerned with the short-term movement of the long-term yield rather than the long-term movement of the short rate. Indeed, even long-term investors must be concerned with short-term fluctuations for risk management purposes, such as value at risk calculations. This paper proposes a new decentralized theory that provides pricing insights for a particular bond or bond portfolio based on the short-term behavior of the yield on that particular bond. The new theory compliments and contrasts with the centralized DTSM approach as it determines the fair value of the yield to maturity on a particular bond based purely on its own near-term dynamics, without making assumptions on how its dynamics will change in the future, or how the instantaneous interest rate or any other bond yields behave.

The theory starts by performing a short-term profit and loss (P&L) attribution to a bond investment through its yield representation. Taking expectation on the P&L attribution under the

risk-neutral measure and setting the expected instantaneous return to the instantaneous interest rate by no-arbitrage leads to a simple pricing equation on the bond yield. The pricing equation decomposes the fair value of the bond yield into three components: its expected rate of change, its risk premium, and the convexity effects driven by its variation. The three components represent three conditional forecasts, but with no reference to where the forecasts come from and how the forecasts vary in the future. The forecasts can come either from estimation of statistical models or directly from expert surveys, the source of which has no bearing on the pricing relation.

Since the yield on each bond can be analyzed on its own, there is no error contagion effect from one security to another. Since the theory only relies on the yield's near-term dynamics, one does not need to make assumptions on how the yield dynamics will change in the future. In particular, when pricing a long-term, the focus is to generate the best conditional mean and variance forecasts on the next movement of its yield, rather than making long-run projections on an unobserved instantaneous interest rate. While the centralized DTSM approach is better suited to perform relative valuation across bonds of different maturities, the new decentralized theory offers a new perspective by analyzing each bond on its own.

When an investor desires to compare a selected basket of bonds, the new theory can also be used for the comparative analysis by directly comparing the risk behaviors of the underlying yields. One can impose common factor structures on the yield changes and generate comparative pricing implications based on the common risk structure. Empirical results from the literature on risk factor analysis of bond yields can thus be readily incorporated into this new pricing framework, making them useful not only for risk management, but also for fair pricing of the bonds under consideration.

As an application, we consider the pricing of long-dated bonds with the assumption of no

directional prediction on its underlying yield. It is extremely difficult to predict the directional movement of long-term yields. We take this difficult-to-predict feature as our starting point, and infer the risk premium component in the long bond yield while controlling for the convexity component based on historical variance estimators on yield changes. We perform empirical analysis on US and UK long-term swap rates and treat them as the coupon rates for par bonds. The analysis shows that the risk premium extracted from each long-term bond can be used for out-of-sample prediction of the future excess returns on the bond.

We also explore the new theory's application to comparative yield curve analysis via common factor structure assumptions on the near-term dynamics. We estimate the convexity effects based on its recent time series behavior, and assume a common factor structure on the yield's rate of change and market price of risk. We propose an estimation framework that extracts the common factors from the observed yield curve and the estimated convexity effects. Estimating the common factor structure on the US and UK swap curves, we show that the extracted common factor on the expected rate of yield change strongly predicts future changes in the yield curve slope.

Finally, butterfly portfolios constructed with three bonds have been a staple trade in the fixed income market. By hedging away systematic interest-rate movements, a well-constructed butterfly portfolio can become very stable, with only small temporal movements left due to the idiosyncratic movements of the particular bonds. Historically, DTSMs have been used to form bond portfolios that are neutral to the common factors of the estimated models (Bali, Heidari, and Wu (2009)). The centralized nature of the theory puts the focus on the factor dynamics assumption while ignoring the particular maturity combination in the butterfly construction. By contrast, the decentralized nature of our new theory allows us to think more about contract characteristics than dynamics assumptions. We propose practical procedures for constructing stable and decentralized butterflies under the new theoretical setting, and show that butterflies constructed with nearby maturities can

be much more stable than butterflies constructed with maturities far apart.

The remainder of the paper is organized as follows. Section 1 develops the new decentralized pricing theory, and contrasts it with the centralized DTSMs. Section 2 describe the data used for the empirical analysis and their summary behaviors. Section 3 examines several practical applications of the new theory. Section 4 concludes.

1. A Decentralized Theory of Bond Yields

We consider an infinite-horizon continuous-time economy. Uncertainty is represented by a filtered probability space $\{\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}\}$, where $\mathbb{P}$ is the physical measure. We assume that the usual conditions of right continuity and completeness with respect to the null sets of $\mathbb{P}$ are satisfied. We further assume the existence of a money market account (MMA) associated with an instantaneous interest rate $r_t \geq 0$. The assumption of no dynamic arbitrage implies the existence of an equivalent martingale measure $\mathbb{Q}$ associated to this MMA numeraire.

Let B_t denote the time-t value of a riskfree bond or bond portfolio with a stream of N cash payments $\{\Pi_j\}_{j=1}^N$ at times $\{t+\tau_j\} \geq t$ for $j = 1, 2, \cdots, N$, with τ_j denoting the time to maturity of the jth payment. Traditional DTSMs start by modeling the dynamics of the pricing kernel or the instantaneous interest rate r_t and value the bond via one of the following expectation operations,

$$
\begin{aligned}
B_t &= \sum_{j=1}^{N} \Pi_j \mathbb{E}_t^{\mathbb{P}} \left[M_{t,t+\tau_j} \right] && (1) \\
&= \sum_{j=1}^{N} \Pi_j \mathbb{E}_t^{\mathbb{P}} \left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) e^{-\int_t^{t+\tau_j} r_u du} \right] && (2) \\
&= \sum_{j=1}^{N} \Pi_j \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^{t+\tau_j} r_u du} \right]. && (3)
\end{aligned}
$$

where $\mathbb{E}_t^{\mathbb{P}}[\cdot]$ and $\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ denote the expectation operator conditional on time-t filtration $\mathcal{F}_t$ under the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$, respectively, $M_{t,T}$ denotes the pricing kernel linking value at time T to value at time t, and $\frac{d\mathbb{Q}}{d\mathbb{P}}$ defines the measure change from $\mathbb{P}$ to $\mathbb{Q}$. The measure change represents the martingale component of the pricing kernel that defines the pricing of various risks. The three equations (1)-(3) represent bond valuation with three different starting points. Through the expectation operation, bonds with cash flows at all times are chained together via (i) the centralized modeling of the pricing kernel dynamics in (1), (ii) the centralized modeling of the instantaneous interest rate dynamics and its market pricing of risk in (2), or (iii) the centralized modeling of the instantaneous interest rate risk-neutral dynamics in (3).

Given the price of a bond B_t, its yield to maturity y_t is defined via the following equality,

$$B_t \equiv \sum_{j=1}^{N} \exp\left(-y_t \tau_j\right) \Pi_j. \tag{4}$$

The yield to maturity can be interpreted as the continuous compounding internal rate of return for holding the bond to expiration.

1.1. Centralized yield decomposition under the classic setting

Before introducing the new decentralized pricing framework, we first perform yield decomposition under the classic centralized setting as a reference. For tractability, we consider the time-t price $B_t(T)$ of a zero-coupon bond that pays \$1 at expiry $T \geq t$. With a single cash flow, the yield to maturity $y_t(T)$ can be explicitly solved from the bond price as,

$$y_t(T) \equiv -\frac{\ln B_t(T)}{T-t}. \tag{5}$$

Substituting the bond pricing formula (3) into the yield equation in (5) reveals the link between the T-expiry yield observed at time $t \in [0,T]$ and future short rates r_u realized at times $u \in [t,T]$:

$$y_t(T) \equiv -\frac{1}{T-t} \ln \mathbb{E}_t^{\mathbb{Q}} e^{-\int_t^T r_u du}, \qquad t \in [0,T]. \tag{6}$$

By adding and subtracting the same term twice, we can decompose the zero-coupon bond yield into three distinct terms,

$$y_t(T) = \mathbb{E}_t^{\mathbb{P}} \frac{\int_t^T r_u du}{T-t} + \mathbb{E}_t^{\mathbb{P}} \left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} - 1 \right) \frac{\int_t^T r_u du}{T-t} \right] - \frac{1}{T-t} \left[\ln \mathbb{E}_t^{\mathbb{Q}} e^{-\int_t^T (r_u - E_t^{\mathbb{Q}} r_u) du} \right]. \tag{7}$$

The first term in the decomposition (7) represents the expectation of the average short rate $\frac{\int_t^T r_u du}{T-t}$ over the life of the bond between now t and the expiry T.

The second term is the risk premium, captured by the covariance under $\mathbb{P}$ of this average short rate with the random variable, $\frac{d\mathbb{Q}}{d\mathbb{P}} - 1$, which has zero mean under $\mathbb{P}$. If interest rates are stochastic and if bond returns are thought to have a positive risk premium, this covariance will also be positive, moving long-term yields above the expected level of the average future short rates.

The third term represents the convexity effect. As the term $C \equiv \frac{1}{T-t} \left[\ln \mathbb{E}_t^{\mathbb{Q}} e^{-\int_t^T (r_u - E_t^{\mathbb{Q}} r_u) du} \right]$ is non-negative, the convexity effect C can only lower the yield. One can interpret C as a *non-standard* deviation under $\mathbb{Q}$ of the zero-mean random variable $-\int_t^T (r_u - E_t^{\mathbb{Q}} r_u) du$. When compared to the standard deviation, the non-standard deviation replaces the quadratic function with an exponential. When the average future short rate $\frac{1}{T-t} \int_t^T r_u du$ is normally distributed under $\mathbb{Q}$ with variance V, the convexity term is equal to $C = \frac{1}{2} V (T-t)$, proportional to the variance of the average future short rate. The positive convexity effect drives long-term yields lower than the expected level of average future short rates.

The relative importance of the three terms in the yield decomposition varies across maturities. As the expiry date T approaches the current time, t, the last two terms vanish, so that the yield-to-maturity approaches the short rate:

$$\lim_{T \downarrow t} y_t(T) = r_t, \qquad t \geq 0. \tag{8}$$

As the time to maturity $\tau = T - t$ increases, the second and the third terms both start affecting the yield, but at different speeds. For illustration on the relative speed of these two effects, consider the very special example where the instantaneous interest rate follows a random walk under $\mathbb{P}$,

$$dr_t = \sigma dW_t \tag{9}$$

and the market price of the bond Brownian risk $(-dW_t)$ is a positive constant $\lambda > 0$.[3] The expected level of future average short rate is simply r_t across all horizons. The risk premium component increases linearly with time to maturity as $\frac{1}{2}\lambda\sigma\tau$, while the convexity effect increases quadratically with time to maturity $\frac{1}{6}\sigma^2\tau^2$. Thus, as the time to maturity increases, first the risk premium component dominates to generate an upward sloping term structure, but ultimately the convexity term starts to dominate and drives the yield curve downward sloping. Researchers often impose mean reversion in the short rate dynamics, which allows both the risk premium and the convexity terms to asymptote to finite constants.

The three-term decomposition of a yield is generic. What is particular about the classic decomposition in (7) is that the three components are linked to the expectation, risk, and pricing of

[3] Given our notation, the market price of the interest rate risk is $-\lambda$. Throughout this paper, even if we start with the interest rate dynamics, we deliberately model the market price of the Brownian risk on the bond, which is the negative of the Brownian risk on the interest rate. This representation aligns better the sign of the market price of risk with that of the bond expected excess return that our analysis focuses on.

one single variable, the instantaneous interest rate, and they are all tied to the expectation operation over the life span of the bond. Thus, to decompose the yield of a long-dated bond under this classic centralized setting, one needs to make projections far into the future about the risk and pricing of the central instantaneous interest rate.

1.2. Decentralized short-run P&L attribution of bond investments

An investor can invest in very long-dated bonds for a very short period of time. In this case, the investor worries more about the short-term value fluctuation than about long-term projections. Even for investors with a long investment horizon, managing the daily P&L fluctuation remains vitally important. Given these practical considerations, our new pricing framework does not rely on long-term projections of a centralized variable (e.g., the instantaneous interest rate), but builds on a decentralized, short-run P&L attribution of the bond investment.

To perform short-run P&L analysis on a bond investment, we focus on the bond value change over the next instant. To decentralize the attribution, we examine how the bond value varies with its own yield to maturity, rather than with the instantaneous interest rate or some other global factors.

We follow industry practice by characterizing the risk of the bond by its duration and convexity, which capture the first- and second-order interest rate sensitivity of the bond value. While there are many variations in the definition, we take the following particular definition that measures the sensitivity of the bond price with respect to its own yield to maturity,

$$\tau \equiv -\frac{\partial B_t}{B_t \partial y_t} = \sum_{j=1}^{N} w_j \tau_j, \quad \tau^2 \equiv \frac{\partial^2 B_t}{B_t \partial y_t^2} = \sum_{j=1}^{N} w_j \tau_j^2, \tag{10}$$

where the weights w_j are given by

$$w_j = \frac{\exp\left(-y_t \tau_j\right) \Pi_j}{\sum_{i=1}^{N} \exp\left(-y_t \tau_i\right) \Pi_i}. \tag{11}$$

According to this definition, the duration (τ) and the convexity (τ^2) are simply the value-weighted average maturity and maturity squared of the cash flows from the bond. The weight on each cash flow is based on its value as a fraction of the total bond worth. For a zero-coupon bond, its duration is simply its time to maturity, and its convexity the maturity squared.

The industry quotes the yield to maturity of a bond instead of its price for stability and comparability across different bonds and over different time periods. The duration and convexity measures capture how much the bond value varies when the yield varies. While both the yield to maturity in (4) and the duration/convexity risk measures in (10) can be compared across different bonds, they are decentralized measures whose calculation depends only on the particular bond itself.

Duration and convexity measures can also be calculated by shocking the yield curve in a particular way. These measures, however, rely on information on the whole yield curve and thus lose the decentralized feature that we desire.

With the decentralized yield to maturity definition and sensitivity measures, we can attribute the short-term P&L of a bond investment with respect to the movement of its own yield to maturity,

$$dB_t = \frac{\partial B_t}{\partial t} dt + \frac{\partial B_t}{\partial y} dy + \frac{1}{2} \frac{\partial^2 B_t}{\partial y^2} (dy)^2 + o(dt), \tag{12}$$

where $o(dt)$ denotes higher-order terms of dt when yield moves diffusively. When the yield can jump randomly, the jump can induce more significant higher-order terms. We henceforth assume that the next move for the yield to maturity of the bond is continuous, and attribute the P&L solely

to time decay and first and second-order effects from the yield to maturity movement. Since the attribution focuses on the bond value change over the next instant, the continuity assumption is only for the next instant. The results hold even if the yield can jump at any other times.

Dividing both sides of equation (12) by $B_t dt$, plugging in the definition of yield in (4) and the definitions of duration and convexity in (10), we attribute of the annualized investment return as,

$$\frac{dB_t}{B_t dt} = y_t - \tau \frac{dy}{dt} + \frac{1}{2}\tau^2 \frac{(dy)^2}{dt}. \tag{13}$$

The first term in (13) denotes the carry. If the bond yield does not change, the instantaneous bond return is simply the yield to maturity. The second term highlights the directional impact of the yield change on the bond return. The negative bond-yield relation dictates that the bond return declines when yield goes up. The sensitivity is measured by the bond's duration τ. The third term captures the convexity of the bond-yield relation. Larger yield moves of either direction increases the bond return due to the convex bond-yield relation. The magnitude of this exposure is captured by the bond's convexity measure τ^2.

Taking expectation on (13) under the statistical measure $\mathbb{P}$, we can attribute the expected bond investment return to three sources,

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{dB_t}{B_t dt}\right] = y_t - \mu_t \tau + \frac{1}{2}\sigma_t^2 \tau^2, \tag{14}$$

where $\mu_t = \mathbb{E}_t^{\mathbb{P}}\left[dy_t/dt\right]$ denotes the time-t expected rate of change on the yield, and $\sigma_t^2 = \mathbb{E}_t\left[(dy)^2/dt\right]$ denotes the time-t conditional variance rate of the yield. Equation (14) decomposes the expected bond return into three sources. The first term captures the expected return from carry. Bonds with a higher yield generate higher returns on average due to carry. Second, due to the negative bond-

yield relation, expected yield increase reduces the expected bond return. Third, higher volatility on the yield movements leads to higher expected bond return due to the convexity effects.

The decomposition highlights the key risk and return sources of bond investments. If an investor has no view on the direction of the yield movement, the investor can form duration-neutral bond portfolios with bonds of nearby maturities by assuming that yields at nearby maturities strongly co-move. The duration-neutral portfolio has minimal exposure to common directional movements of the bond yields, and the long-short positioning of the bonds in the portfolio can be driven by the difference between the carry and convexity benefits of the bonds in the portfolio.

To illustrate the contributions from the different components, let us imagine a situation where zero-coupon bond yields at long maturities (e.g., 10, 15, 30) are identical and move in parallel by substantial amount (i.e., σ^2 is large). Under this imaginary situation, one can form a self-financing and riskless bond portfolio and make money from long convexity. First, since the yields are the same, a dollar-neutral portfolio would be self-financing. Second, since the yields move in parallel, a duration-neutral portfolio will have no directional exposure. If such a portfolio can be formed with positive convexity, one would expect to market positive money in the future. As a concrete example, we can form a butterfly portfolio by being long \$300 10-year zero-coupon bond, long \$100 30-year zero-coupon bond, and short \$400 15-year zero-coupon bond. The butterfly portfolio costs zero dollar (dollar-neutral), has zero duration, and contains a positive dollar convexity of $c_f = 75$. The instantaneous P&L on the butterfly is nonnegative and proportional to its dollar convexity and the squared yield change,

$$df_t = \frac{1}{2} c_f (dy)^2 \geq 0. \tag{15}$$

Therefore, observing a flat and parallel moving yield curve presents an arbitrage opportunity.

1.3. Decentralized no-arbitrage pricing and yield decomposition

The P&L attribution analysis highlights the local risk sources and return opportunities for the bond investment over the next instant. To generate pricing implications, we take expectation under the risk-neutral measure $\mathbb{Q}$ on the attribution in (13),

$$\mathbb{E}_t^{\mathbb{Q}}\left[\frac{dB_t}{B_t dt}\right] \quad = \quad y_t - \mu_t^{\mathbb{Q}}\tau + \frac{1}{2}\sigma_t^2\tau^2, \tag{16}$$

where $\mu_t^{\mathbb{Q}} = \mathbb{E}_t^{\mathbb{Q}}[dy_t/dt]$ denotes the time-t expected rate of yield change under the risk-neutral measure. Given the diffusive assumption over the next instant with the instantaneous volatility σ_t, if we use λ_t to denote the market pricing of the bond Brownian risk, we can link the expected rate of yield change under the two measures by

$$\mu_t^{\mathbb{Q}} = \mu_t + \lambda_t\sigma_t. \tag{17}$$

The absence of dynamic arbitrage dictates that the risk-neutral expected instantaneous rate of return on any investment is equal to the instantaneous interest rate r_t. Applying this no-dynamic-arbitrage condition to the risk-neutral expectation in (16) leads to a simple pricing relation for the bond yield spread over the instantaneous interest rate,

$$y_t - r_t = \mu_t\tau + \lambda_t\sigma_t\tau - \frac{1}{2}\sigma_t^2\tau^2. \tag{18}$$

The fair value of the yield spread $(y_t - r_t)$ on the bond investment is determined by its expected rate of change forecast (μ_t), its risk premium $(\lambda_t\sigma_t)$, and its volatility forecast (σ_t).

Theorem 1 *If the yield of a bond is moving continuously over the next instant, no dynamic arbi-*

trage dictates that the fair spread of this yield over the instantaneous interest rate is linked to its expected rate of change (μ_t), its risk premium ($\lambda_t\sigma_t$), and its variance rate (σ_t^2) through the bond's duration τ and convexity τ^2 by

$$y_t - r_t = \mu_t\tau + \lambda_t\sigma_t\tau - \frac{1}{2}\sigma_t^2\tau^2. \tag{19}$$

Compared to classic centralized bond pricing, the new pricing relation in (19) is highly decentralized. The fair valuation of the bond investment in (19) only depends on the behavior of its own yield to maturity, with no direct dependence on the short rate dynamics or the dynamics of any other yields. In fact, the pricing does not even rely on the full dynamics of its own yield, but only depends on the conditional expectation estimators of its rate of change, its volatility, and its market pricing. All three estimators can change over time, so can the dynamics of the yield, but none of these changes enter the pricing of the current yield spread. One can bring in the forecasts from any outside sources and directly examine the pricing implication under the new theory. These forecasts can come from any model assumptions, algorithms, or expert opinions, allowing maximum flexibility and cross-field collaboration.

The pricing relation in (19) also provides a decentralized version of the yield decomposition. Similar to the centralized yield decomposition, equation (19) also decomposes the yield into expectation, risk premium, and convexity. The difference is that the expectation, risk premium, and convexity in (19) are all measured on the yield of this particular bond. They reflect the expected behavior of the yield over the next instant, rather than the behavior of the short rate over the whole life span of the bond.

To distinguish our new decentralized pricing theory from the centralized DSTMs, we henceforth label our new pricing theory as **Dynamic Decentralized Yield Model** (DDYM).

1.4. Comparison with DTSMs

The new theory links the time-t value of the yield spread of a particular bond to the current forecasts of its rate of change, its risk premium, and its variance rate, thus making the analysis completely decentralized to the present and on the particular bond. The decentralized feature also dictates that the no-arbitrage relation in (19) only guarantees dynamic no-arbitrage at one point in time between the particular bond in consideration and the money market account given the three conditional forecasts. It has no direct implications on the cross-sectional relations across yields on different bonds, nor on the pricing of bonds at any other points in time.

By comparison, DTSMs derive the fair value for the whole yield curve based on the full dynamics specification of a centralized instantaneous interest rate, and thus guarantee dynamic consistency across the whole yield curve. While the new theory allows one to compare the level of a yield to its own conditional moment forecasts, DTSMs are built to analyze the yield curve shape and to make cross-sectional comparisons of yields across different maturities. The two theories are complimentary in focusing on different aspects of dynamic no-arbitrage.

1.4.1. Decentralizing DTSM to DDYM

Since DTSMs are derived based on dynamic no arbitrage of all yields relative to one centralized instantaneous interest rate process, the level of each derived yield is naturally consistent with the derived dynamics of this yield. Thus, the level of the derived yield and the levels of the derived rate of change and volatility of the yield must satisfy our DDYM pricing relation.

As an example, consider the general diffusion-affine dynamic term structure model of Duffie and Kan (1996), who assume that the instantaneous interest rate is an affine function of K factors

with affine continuous dynamics under the risk-neutral measure:

$$r_t = a_r + b_r^\top X_t \tag{20}$$

$$dX_t = \kappa(\theta - X_t)\,dt + \sqrt{\Sigma(X_t)}dZ_t, \tag{21}$$

with $\Sigma(X_t)_{ii} = \alpha_i + \beta_i^\top X_t$ and $\Sigma_{ij} = 0$ for $i \neq j$. Under these dynamics assumptions, the yields on all zero-coupon bonds are affine in the K factors,

$$y_t(T) = \frac{a(\tau)}{\tau} + \left[\frac{b(\tau)}{\tau}\right]^\top X_t, \tag{22}$$

for all $\tau = T - t > 0$, where the coefficients $(a(\tau), b(\tau))$ are solutions to the following set of ordinary differential equations,

$$a'(\tau) = a_r + b(\tau)^\top \kappa\theta - \frac{1}{2}\sum_i b(\tau)_i^2\,\alpha_i, \tag{23}$$

$$b'(\tau) = b_r - \kappa^\top b(\tau) - \frac{1}{2}\sum_i b(\tau)_i^2\,\beta_i, \tag{24}$$

starting at $a(0) = 0$ and $b(0) = 0$.

Equation (22) centralizes the yields of all maturities by linking them as affine functions of a common set of factors X_t. To decentralize the affine model, we take one particular zero-coupon bond with an expiry T as an example and derive the risk-neutral dynamics of the yield on this zero-coupon bond. Applying Ito's lemma to (22) and (21), we have the risk-neutral rate of change and variance of the yield $y_t(T)$ as

$$\begin{aligned}
\mu_t^{\mathbb{Q}} &= -\left[\frac{a'(\tau)}{\tau} - \frac{a(\tau)}{\tau^2}\right] - \left[\frac{b'(\tau)}{\tau} - \frac{b(\tau)}{\tau^2}\right]^\top X_t + \left[\frac{b(\tau)}{\tau}\right]^\top \kappa(\theta - X_t), \\
\sigma_t^2 &= b(\tau)^\top \Sigma(X)\, b(\tau)\,\frac{1}{\tau^2}.
\end{aligned}$$

Applying the DDYM relation in (19) and the instantaneous interest rate function in (20),

$$\begin{aligned} y_t(T) &= r_t + \mu_t^{\mathbb{Q}}\tau - \frac{1}{2}\sigma_t^2\tau^2 \\ &= a_r + b_r^\top X_t - a'(\tau) + \frac{a(\tau)}{\tau} + b(\tau)^\top \kappa\theta - \left[b'(\tau) - \frac{b(\tau)}{\tau} + b(\tau)\kappa\right]^\top X_t \\ &\quad -\frac{1}{2}\sum_i b(\tau)_i^2 \alpha_i - \frac{1}{2}\sum_i b(\tau)_i^2 \beta_i^\top X_t, \end{aligned}$$

which is affine in X_t. Applying (22) and collecting terms, we have

$$\begin{aligned} \frac{a(\tau)}{\tau} &= a_r - a'(\tau) + \frac{a(\tau)}{\tau} + b(\tau)^\top \kappa\theta - \frac{1}{2}\sum_i b(\tau)_i^2 \alpha_i, \\ \frac{b(\tau)}{\tau} &= b_r - \left[b'(\tau) - \frac{b(\tau)}{\tau} + b(\tau)\kappa\right] - \frac{1}{2}\sum_i b(\tau)_i^2 \beta_i. \end{aligned}$$

Rearrange, the DDYM pricing relation leads to the same ordinary differential equation as the affine model in (23)-(24),

$$\begin{aligned} a'(\tau) &= a_r + b(\tau)^\top \kappa\theta - \frac{1}{2}\sum_i b(\tau)_i^2 \alpha_i, \\ b'(\tau) &= b_r - b(\tau)\kappa - \frac{1}{2}\sum_i b(\tau)_i^2 \beta_i. \end{aligned}$$

The DDYM pricing relation starts with $\mu_t^{\mathbb{Q}}$ and σ_t^2 and derive their linkage to the yield level via dynamic no-arbitrage arguments, without referring to other parts of the curve. The DTSM starts with the instantaneous interest rate dynamics and derives the whole yield curve by taking expectations over future payoffs of each bond, with the expectation operation formulated based on the instantaneous interest rate dynamics. The derived yield level and the derived yield dynamics naturally satisfy the no-arbitrage relation that we derive. By deriving everything from a centralized instantaneous interest rate dynamics, DTSM allows one to compare the cross-sectional behavior

of whole yield curve. By focusing on the conditional moment conditions of itself, DDYM allows one to link the level of one particular yield to its own risk, expectation, and pricing.

1.4.2. Centralizing DDYM to DTSM

The DDYM approach derives the no-arbitrage relation between the level of one particular yield and its near-term dynamics. By imposing a functional linkage on the near-term dynamics across the whole yield curve, we can in principle centralize the DDYM relation to arrive something close to a dynamic term structure model. This centralization process, however, is not always straightforward because it is not a trivial task to simultaneously assume the near-term dynamics of all yields without introducing arbitrages among them. In what follows, we show one particularly simple example that allows us to do the centralization without introducing cross-sectional arbitrage.

Assume that the continuously compounding yield curve goes up and down in parallel under the physical measure $\mathbb{P}$. If we use $y_t(\tau)$ to denote the yield at a fixed time to maturity τ, we can write its dynamics as,

$$dy_t(\tau) = \sigma dW_t^{\mathbb{P}}, \tag{25}$$

for all $\tau \geq 0$, where we assume zero drift and the same volatility σ for all maturities τ.

Further assume that the market price of the bond Brownian risk $(-dW_t)$ is λ, we can derive the yield dynamics under risk-neutral measure $\mathbb{Q}$ as,

$$dy_t(\tau) = \lambda\sigma dt + \sigma dW_t. \tag{26}$$

Consider a zero-coupon bond with fixed expiry T. The parallel shifting yield curve implies that

the risk-neutral dynamics for the yield of this zero-coupon bond $y_t(T)$ can be written as,

$$dy_t(T) = dy_t(\tau) - y_t'(\tau)\,dt = \left[\lambda\sigma - y_t'(\tau)\right]dt + \sigma dW_t, \tag{27}$$

where the $y'(\tau)$ term accounts for the sliding of the yield on this bond along the yield curve.

Starting with the drift and diffusion in (27), we can apply our DDYM pricing relation in (19), and represent the yield as,

$$y_t(\tau) = r_t + \left(\lambda\sigma - y_t'(\tau)\right)\tau - \frac{1}{2}\sigma^2\tau^2, \tag{28}$$

for all τ.

Define $z_t(\tau) \equiv y_t(\tau)\tau = -\ln B_t(T)$, equation (28) implies

$$z_t'(\tau) = r_t + \lambda\sigma\tau - \frac{1}{2}\sigma^2\tau^2, \tag{29}$$

for all τ. Thus, under this particular parallel shift assumption and taking logs on the zero-coupon bond price curve, the short rate is simply the negative of the slope of this log price curve at zero maturity, the risk premium ($\lambda\sigma$) is the curvature of this log price at zero maturity, and the yield variance σ^2 is the third derivative of the log price curve.

We can solve for the whole yield curve by integrating equation (29) over maturity,

$$y_t(\tau) = \frac{z(\tau)}{\tau} = \frac{1}{\tau}\int_0^\tau \left(r_t - \lambda\sigma u - \frac{1}{2}\sigma_t^2 u^2\right)du = r_t - \frac{1}{2}\lambda\sigma\tau - \frac{1}{6}\sigma^2\tau^2, \tag{30}$$

for all τ. By assuming parallel shift on the yield curve and by specifying the full dynamics of all yields, equation (30) centralizes the DDYM relation in (28) to arrive at a term structure model.

If we only assume the near-term dynamics by allowing σ_t and λ_t to vary over time with unknown dynamics, the differential equation in (28) remains valid, but we can no longer perform the integration in (30) without knowing the full path of σ_t and λ_t from t to T. Deriving the full term structure model necessitates the specification of the full dynamics.

Merton (1973) considers in a footnote a similar model with $dr = \sigma dW_t^{\mathbb{P}}$ and arrives at a similar term structure that excludes arbitrage. This particular example not only guarantees no arbitrage between the particular zero-coupon bond and the instantaneous rate, but also guarantees that bonds across all finite maturities do not allow arbitrage. To verify, we can start with $dr = \lambda\sigma dt + \sigma dW_t$. Then the relation in (30) between $y_t(\tau)$ and r_t suggest that $dy_t(\tau) = dr_t = \lambda\sigma dt + \sigma dW_t$, just as we have assumed to begin with.

2. Data and Summary Statistics

We perform empirical analysis using US and UK swap rates. The financing leg of the swap contracts for both currencies are the 6-month LIBOR rate of the corresponding currency. We can treat the swap rates as the coupon of a par bond. We obtain the LIBOR and swap rate data from Bloomberg, daily from January 4, 1995 to December 29, 2017, spanning 5,790 business days. The swap maturities include 2, 3, 4, 5, 7, 10, 20, 30, 40, and 50 years. Swap rates at short maturities are available over the whole sample period. Longer-term swap rates become available at a later date. For the US, the 40- and 50-year swap rates become available starting November 12, 2004. For the UK, the starting dates for the 15-, 20-, 30-, 40, and 50-year swap rates are April 10, 1997, June 2, 1997, January 19, 1999, September 7, 1999, and August 8, 2003, respectively. When computing summary statistics and estimating models, we sample the data weekly every Wednesday from January 4th, 1996 to December 27, 2017, for 1,200 weeks. The weekly sampling is to avoid week-day

effects, and we leave out the first year of the sample for constructing historical variance estimators.

Table 1 reports the summary statistics of the swap rate series (in percentage points), including the sample average ("Mean"), standard deviation ("Stdev"), skewness ("Skew"), excess kurtosis ("Kurt"), and weekly autocorrelation ("Auto"). Panel A reports the statistics on the swap rate levels. The mean swap rate term structure is initially upward sloping for both economies, but the mean estimates become lower at very long maturities. The standard deviation of the swap rate series show a declining term structure, suggesting that longer rates vary within a narrower range. For both economies, the skewness estimates are small and the excess kurtosis estimates are negative. The weekly autocorrelation estimates for all swap series are very close to one (0.994 to 0.999), suggesting that the swap series are highly persistent, if at all stationary.

[Table 1 about here.]

Panel B of Table 1 reports the summary statistics on the weekly changes of the swap rates. The mean and standard deviation estimates of the weekly changes are annualized. The annualized sample averages of the weekly changes are negative for all series, suggesting that interest rates have been showing a declining trend over the past two decades for both economies.

The annualized standard deviation estimates of the weekly swap rate changes represent the unconditional volatility estimators for the interest rate changes. The estimators show remarkable stability across maturities. For the US swap rates, the volatility estimates range from 0.82 at two-year maturity to 0.95 at seven-year maturity. The volatility estimates for swap rates from 20 to 50 years are virtually the same around 0.86 of a percentage point. For the UK, the volatility estimates range from 0.60 at 50-year maturity to 0.79 at three- and four-year maturities, showing again a very flat volatility term structure. The skewness of the weekly changes remain small, but the excess kurtosis estimates of weekly changes are all positive. The autocorrelation estimates on the

weekly changes are all close to zero.

Figure 1 plots the time series of the swap rates at three selected maturities: 2-year (solid lines), 10-year (dashed lines), and 30-year (dash-dotted lines), for the US in Panel A and the UK in Panel B. In line with the negative mean swap rate change estimates in Panel B of Table 1, the swap rates show a distinct downward trend for both economies during the 20-plus year sample period, reflecting the downward trend in inflation during this sample period. We overlay the swap rate time series with the recession band for each economy. During our sample period, the US economy has experienced two recessions, a minor one in 2001 and the more severe one often dubbed as the great recession in 2008-2009. The UK economy did not experience a recession in 2001, but shared the great recession in 2008-2009. At the start of a recession, the central bank tends to start cutting the short-term interest rate in an effort to stimulate the economy. Such actions are reflected in the sharp short rate drops during the shaded recession period. The US economy recovered quickly after the 2001 minor recession. The short-term swap rate also went up quickly after the initial drop. After the great recession, however, the short-term rate stayed low for a much longer period.

[Figure 1 about here.]

Based on the daily changes of the swap rates series, we construct annualized volatility estimators on each series with a one-year rolling window. Figure 2 plots the time series of the rolling volatility estimates at selected maturities, overlaid with the recession bands. For each swap rate series, the rolling volatility estimates vary strongly over time, reaching its peak during the 2009 financial crises but having been calming down since then. Across maturities, the volatility estimates show both upward and downward sloping term structure patterns. The term structure tends to be downward sloping when the volatility level is high, and upward sloping during more quiet periods.

[Figure 2 about here.]

The volatility estimates tend to be high during transition periods in a business cycle when the central bank is actively cutting or raising rates. During these periods, the volatility estimates for short-term rates tend to be higher than the estimates for long-term rates, leading to downward sloping volatility term structures. For the US, the Fed was very active in 2001 as it cut the Fed fund rate target 11 consecutive times. It was actively raising rates from June 2004 to June 2006 for 17 consecutive times. During the beginning of the great recession, the Fed cut the Fed fund rate aggressively from 4.25% in December 2017 to virtually 0 by December 2008. During these periods, the volatility estimators for the short-term rates are higher than for the long-term rates.

On the other hand, during periods with few central bank actions on the short rate, such as during the past eight years since 2010 as the short rate was trapped at virtually zero, the volatility term structure becomes distinctively upward sloping. The upward sloping volatility term structure is interesting and can prove challenging for classic models that assume mean-reverting dynamics on the instantaneous interest rate. The new pricing theory has no difficulty accommodating the upward sloping volatility term structure and the large volatility estimates for long-dated interest rates. The theory allows one to directly take the volatility estimators as inputs without making explicit assumptions on the short rate dynamics.

3. Applications

We explore practical applications of the new pricing theory from several angles. First, we propose to predict future bond excess returns while assuming no predictability on long-dated floating interest rate series. Second, we perform comparative analysis of the yield curve via common factor structure assumptions. Third, we re-examine the well-known butterfly trades under our new pricing framework based on local commonality assumptions.

3.1. Predicting long bond returns with no rate predictability

It is extremely difficult to predict long-term interest rate movements. The expectation hypothesis literature proposes to predict future short-rate changes based on the slope of the yield curve, but it does not provide much help in predicting movements of long rates. The DTSM literature often assumes mean-reverting dynamics for the short rate, but the assumption often implies overly low volatility for long rates. In this section, as an application of the new pricing theory, we take no-predictability on long-term rates as the starting point, and infer bond risk premium from the observed interest rate level and interest rate volatility estimators.

We start by assuming that the constant-maturity floating yield $y_t(\tau)$ at some long fixed time to maturity τ moves diffusively like a random walk over the next instant,

$$dy_t(\tau) = \sigma_t(\tau) dW_t^{\mathbb{P}}, \tag{31}$$

where $\sigma_t(\tau)$ denotes the time-t conditional forecast of the volatility rate of this yield. The conditional volatility can vary over time with unspecified dynamics. It is also possible that the volatility forecasts differ for yields of different maturities. The key assumption underlying (31) is the absence of predictability on the long-term floating yield as the time-t expected rate of change is assumed to be zero.

If we denote the market price of the risk for the corresponding bond $(-dW_t)$ as λ_t, we can derive the risk-neutral dynamics for the yield as

$$dy_t(\tau) = \lambda_t \sigma_t(\tau) dt + \sigma_t(\tau) dW_t. \tag{32}$$

The risk-neutral drift for the fixed-expiry yield $y_t(T)$ is further adjusted by the local shape of the

yield curve as it slides along the yield curve,

$$\mu_t^{\mathbb{Q}} = \lambda_t \sigma_t(\tau) - y_t'(\tau). \tag{33}$$

Plugging the no-prediction dynamics into the DDYM pricing relation in (19), we have

$$y_t = r_t + \lambda_t \sigma_t(\tau)\tau - y_t'(\tau)\tau - \frac{1}{2}\sigma_t^2(\tau)\tau^2. \tag{34}$$

Rearrange, we have,

$$y_t + y'(\tau)\tau = r_t + \lambda_t \sigma_t(\tau)\tau - \frac{1}{2}\sigma_t^2(\tau)\tau^2. \tag{35}$$

For zero-coupon bonds, $y_t + y'(\tau)\tau = \frac{\partial(y_t\tau)}{\partial\tau} = f_t(\tau)$ is the instantaneous forward rate. For coupon bonds, we can directly estimate the yield curve slope against the bond duration based on observed yields at nearby maturities and treat the slope-adjusted yield $y_t + y'(\tau)\tau$ as a forward yield analog.

Equation (35) shows that in the absence of rate prediction, positive market price of bond risk drives the yield curve up with increasing duration whereas convexity drives the curve down. Based on observed yield curve time series, we can construct volatility estimators for yield changes across different maturities to generate a volatility term structure curve $\sigma_t(\tau)$ at each date. We can also use the observed yield curve to infer the interest rate level and slope at the corresponding maturity. Combining these observations and estimates with the financing cost r_t, we can infer the market price of bond risk.

Proposition 1. *Under the assumption of no-predictability on long-term constant-maturity yields as in (31), the market price of the bond risk can be estimated from the observed yield curve shape*

and the volatility estimator of the yield changes,

$$\lambda_t = \frac{y_t(\tau) + y'(\tau)\tau - r_t + \frac{1}{2}\sigma_t^2(\tau)\tau^2}{\sigma_t(\tau)\tau}. \tag{36}$$

For empirical implementation of the proposition, we assume that long-dated US and UK swap rates with maturities 10 years and longer are not predictable. We take the financing rate (6-month LIBOR) as the short rate r_t, construct volatility estimators on daily changes of each swap rate series with a one-year rolling window, and treat the swap rates as par bond yields in estimating the duration and convexity of these par bonds. Finally, we estimate the swap rate curve slope against the duration at each maturity using a local linear regression.

Without rate prediction, Equation (36) sets the bond risk premium as the forward yield spread adjusted for the convexity contribution, $C_t = \frac{1}{2}\sigma_t^2(\tau)\tau^2$. Figure 3 plots the time series of the estimated convexity contribution C_t at selected maturities. The convexity contribution is negligible at short maturities, but becomes significant at long maturities. In the US, the convexity contribution estimates are over 100 basis points for 30- and 50-year swap rates during the volatile period of 2009. In the UK, the swap rate volatility estimates are lower. The convexity contribution for 30- and 50-years swaps varies between 20 and 40 basis points.

[Figure 3 about here.]

Figure 4 plots the time series of the extracted market price of bond risk (λ_t) from each swap series from 10 to 50 years. The market price of bond risk extracted from different swap rate series are similar in magnitude and move closely together. Over the common sample, the cross-correlation estimates among the different λ_t series average 99.44% for the US swap rates and 99.48% for the UK swap rates. Similar to findings in Cochrane and Piazzesi (2005), our evidence

supports a one-factor structure for the bond risk premium.

[Figure 4 about here.]

On average, the market price of bond risk estimates are positive for both economies, supporting the hypothesis of positive average bond risk premium. Nevertheless, the estimates vary strongly over time. In the US, the estimates become close to zero right before the start of the two recessions in 2000 and 2007, respectively, but the estimates become the most positive at the end of each recession. In the UK, the market price of bond risk went into negative territory in 1998 and again between 2007 and 2008, but otherwise show positive co-movements with the US economy.

To examine whether the ex-ante risk premium estimate ($\lambda_t \sigma_t \tau$) on each long-term swap rate series predicts future excess returns on the corresponding par bond, we compute six-month and one-year ahead excess returns on the par bonds and measure the forecasting correlation between the ex ante risk premium and ex post excess returns on each par bond. Table 2 reports the forecasting correlation estimates. The ex ante risk premium estimates show strongly positive predictive power on the ex post bond return. At six-month horizon, the forecasting correlation estimates for the US range from 0.23 to 0.31. The estimates for the UK range from 0.17 to 0.21. At one-year horizon, the forecasting correlation estimates become higher and more uniform, ranging from 0.30 to 0.37 for the US and 0.31 to 0.35 for the UK.

[Table 2 about here.]

The assumption of no prediction on long-dated swap rates leads to significant prediction on bond excess returns. The forecasting correlation estimates are quite high, particularly when recognizing that the forecasting correlation estimates are fully out of sample. The risk premium estimates at time t are constructed based on the observed yield curve at that time and the yield curve's

volatility over the past year, without any further calibration or forecasting regression involved.

In related literature, Cochrane and Piazzesi (2005) propose to predict bond excess returns with a portfolio of forward rates. They find that the excess returns on Treasury bonds across different maturities can be predicted by the same forward rate portfolio, similar to our finding of one dominant factor for the market price of bond risk. Cieslak and Povala (2015) propose to decompose Treasury yields into long-horizon inflation expectations and maturity related cycles, and show that, with the long-run inflation expectation removed, a portfolio of the extracted cycles can predict the bond excess returns better. Several studies also link the bond risk premia to macroeconomic indicators of business cycle variations.[4]

The DDYM model decomposes the bond yields into the expectation of the yield changes, risk premium, and convexity effects. While the convexity effects can be estimated based on historical volatility estimators, separating the risk premium from the expectation remains a challenging task. Kim and Orphanides (2012) propose to use survey data to pin down interest rate expectation for better identification of the bond risk premium in a DTSM setting. Similar survey data can readily incorporated into our DDYM model for risk premium identification. In this section, by focusing on the long segment of the yield curve where predicting interest rate changes is inherently difficult, we turn the lack of interest-rate predictability on its head and use the no-predictability as an effective identification condition for the bond risk premium.

3.2. Comparative yield analysis via common factor structures

The new theory is built to analyze the relation between the yield of a particular bond and its own near-term dynamics. To perform comparative yield analysis across different maturities, one can

[4]See, for example, Fama and French (1989), Cooper and Priestley (2009), Ludvigson and Ng (2009), and Gargano, Pettenuzzo, and Timmermann (2017).

assume common factor structures in their near-term dynamics and derive the implications of the common factor assumptions on the yield curve. For illustration, this section considers a particularly simple common factor structure for the swap rate curve dynamics and explore its implications.

3.2.1. Common factor structures along the term structure

We make the following common factor structure assumptions on the expected rate of change, market price, and volatility of the swap rates across different maturities:

Assumption 1 (Expected rate of change term structure). *The expected rate of change on the constant-maturity yields vary across duration via an exponential form,*

$$\mu_t(\tau) = e^{-\kappa_t \tau}(\mu_{t,0} - \mu_{t,\infty}) + \mu_{t,\infty}. \tag{37}$$

Under the assumption, the expected rate of change converges to μ_t^0 as duration approaches zero and converges to μ_t^∞ as the duration approaches infinity. The exponential form is a particularly simple way of capturing the gradual transition from one limit to another.

Assumption 2 (No predictability on long rates). *The expected rate of change approaches zero as the duration approaches infinity:*

$$\mu_{t,\infty} = 0. \tag{38}$$

We maintain the no-predictability assumption in the long-run limit.

Assumption 3 (Market price of risk). *The market price of bond risk is identical across maturities,*

$$\lambda_t(\tau) = \lambda_t. \tag{39}$$

The one-factor structure on the market price of risk is in line with the empirical evidence from the previous section and the literature findings.

Assumption 4 (Volatility). *The annualized volatility rates of yield changes are equal to their corresponding historical estimators* $\widehat{\sigma}_t(\tau)$.

Out of the three components of the yield decomposition, the convexity effect can be constructed based on historical volatility estimators, and the market price of bond risks has been found to share a one-factor structure. This leaves the yield change expectation as the most variable and hard-to-identify component. The previous section makes the identification by focusing on the long end of the yield curve and assuming no-predictability on the long rates. This section allows non-zero expected rate of change in the short end of the yield curve, and assumes that the expected rate of exchange varies smoothly across maturity, controlled by an exponentially decaying function.

With assumptions (1) to (4), we can write the term structure of the yield curve as

$$o_t(\tau) = y_t(\tau) + y'(\tau)\tau - r_t + \frac{1}{2}\widehat{\sigma}_t^2(\tau)\tau^2 = e^{-\kappa_t\tau}\mu_{t,0}\tau + \lambda_t\widehat{\sigma}_t(\tau)\tau + e_t, \tag{40}$$

where we move the observable quantities to the left hand side of the equation and leave the parametric component of the yield curve to the right hand side. The yield level $y_t(\tau)$ and the financing rate r_t are directly observable. We estimate the yield curve slope $y_t'(\tau)$ using a local linear regression, and we construct the volatility estimator $\widehat{\sigma}_t(\tau)$ using one-year of historical daily yield changes. We label the yield spread adjusted by the curve slope and convexity as $o_t(\tau)$ and treat it as an observable quantity, potentially with noise e_t.

The right hand side of equation (40) includes the parametric specification on the expected rate of change and the market pricing of bond risk. At each date t, the specification governs the yield

curve term structure via three variables $(\mu_{t,0}, \kappa_t, \lambda_t)$. Intuitively, a positive market price of bond risk λ_t contributes to a positive slope to the term structure. The expected rate of change on the short end $\mu_{t,0}$ further adjusts the slope through the expectation difference across maturities, and the speed of decay κ_t controls the curvature of the slope and the speed by which the expectation contribution declines as maturity increases.

3.2.2. Identifying the common factors from the observed yield curve

With the common factor structure assumptions, the yield curve at any given date t is governed by three variables $(\mu_{t,0}, \kappa_t, \lambda_t)$. One particular feature of the new theory is that the yield curve at time t depends on the levels of these variables at time t, but does not depend on the particular dynamics specification for these variables. Therefore, the emphasis of the empirical analysis involves the extraction of the state variables from the yield curve, without knowing the state dynamics.

Based on this unique feature, we cast the model into a state-space form, where we treat the three state variables as the hidden states and treat the observed yield curve o_t as measurements with errors. We define the state vector as X_t,

$$X_t \equiv [\mu_{t,0}, \ln(\kappa_t), \lambda_t]^\top, \tag{41}$$

where the logarithm transform on κ_t guarantees that it stays strictly positive and the state vector can take values on the whole real line. Since how the state variables vary over time does not affect the pricing, we can specify the state propagation equation without worrying about their pricing implications. We make the particularly simple assumption of a random walk dynamics,

$$X_t = X_{t-1} + \sqrt{\Sigma_x}\varepsilon_t. \tag{42}$$

where the standardized error vector ε_t is assumed to be normally distributed with zero mean and unit variance. We further assume that the covariance matrix is a diagonal matrix with distinct diagonal values so that the states can have different degrees of variation but the movements are independent of each other.

We define the measurement equations on the observed yield spread,

$$o_t = h(X_t, \tau) + e_t, \tag{43}$$

where o_t denotes the ten yield spread series and $h(X_t)$ denotes the value of the yield spread as a function of the states X_t and yield maturity τ, as defined by the specification in equation (40), and e_t denotes the measurement error on the observed yield. We assume that the pricing errors are iid normally distributed with error variance σ_e^2.

When the state-space model is Gaussian linear, the Kalman (1960) filter provides efficient forecasts and updates on the mean and covariance of the state and observations. Our state-propagation equations are constructed to be Gaussian and linear, but the measurement functions $h(X_t)$ are not linear in the state vector. We use the unscented Kalman filter (Wan and van der Merwe (2001)) to handle the nonlinearity.

The setup introduces four auxiliary parameters that define the covariance matrices of the state propagation errors and the measurement errors. The relative magnitude of the state propagation error variance versus the measurement error variance controls the speed with which we update the states based on new observations. Intuitively, if the states vary a lot (large Σ_x) and the observations are accurate (small σ_e^2), one would want to update the states faster to better match the new observations. If on the other hand the states vary slowly over time and the observations are very noisy, one would want to update the states more slowly to smooth out the noise in the observation. We

choose the magnitudes for these auxiliary parameters, and accordingly the optimal state updating speed, by minimizing the sum of squared forecasting errors in a quasi maximum likelihood setting.

3.2.3. Pricing performance of the common factor structure

Table 3 reports the summary statistics of the model's pricing errors, defined as the basis point difference between the observed swap rates and the model-generated values. The statistics include the sample average of the pricing error ("Mean"), root mean squared pricing error ("RMSE"), weekly autocorrelation ("Auto"), and the model's explained variation ("EV") on each series, defined as one minus the variance ratio of the model's pricing error to the original volatility series. For both economies, the mean pricing errors are positive at intermediate maturities from 10 to 20 years, but negative at both short and very long maturities, suggesting that the assumed factor structure is not flexible enough to fully capture the curvature of the swap rate curve. Inspection of the pricing error time series reveals that the simple factor structure has difficulties capturing the strong curvature of the swap curve during an extended period of our sample when the short rate hits the lower bound. When the short rate hits the lower bound, the shadow rate can be very negative[5] and the observed term structure for the expected rate of change is more S-shaped than exponential.

[Table 3 about here.]

The root mean squared pricing errors average around 10 basis points. The explained variation is somewhat lower at short maturities, but over 99% for maturities at five years and longer. The autocorrelation estimates for the pricing errors are much smaller than that for the original swap series reported in Table 1, suggesting that the factor structure is reasonably successful in separating

[5] Black (1995) proposes the idea of shadow rate and treats observed nominal interest rate as options on the shadow rate. Krippner(2012, 2013), among others, proposes tractable implementations of shadow rate models. Christensen and Rudebusch (2015) estimate a three-factor shadow rate on Japanese yields.

the systematic common movements on the yield curve from temporal dislocations at particular maturities. The highest autocorrelation estimates come from 15- and 50-year maturities, coinciding with the highest root mean squared pricing error.

3.2.4. Predicting changes in yield curve slope with the extracted rate of change

Figure 5 plots the time series of the common market price of bond risk λ in the solid line. For comparison, we have also overlaid in dashed lines the market price of risk estimates from the previous subsection based on no-predictability assumption on long-term rates. The two sets of estimates follow each other closely, with cross-correlation estimates ranging from 75% to 84%. For the US, the two sets of estimates are virtually identical before the 2009 great recession. Since the recession, as the short-term rates hit the zero lower bound, and the swap rate curve shape becomes more complicated than can be captured by a simple exponential function, and the two sets of estimates start to show some deviations. For the UK, the largest deviation between the two sets of estimates also happen right after the 2009 recession.

[Figure 5 about here.]

Figure 6 plots the time series of the extracted state variable μ_0, which measures the relative expected rate of change at short maturities versus long maturities. The expected annualized rate of change varies within a band of $(-1.84\%, 1.82\%)$ for the US and within a slightly narrower range $(-1.72\%, 1.23\%)$ for the UK. The time series show large variations following the business cycle, as well as shorter-term temporal variations. For example, for the US, the expected rate of change switched signs several times between 1996 to 2006.

[Figure 6 about here.]

Figure 7 plots the time series of the reciprocal of the decay estimates, $1/\kappa$. κ measures the speed at which the expected rate of change exponentially declines with maturity. The reciprocal of κ provides an intuitive time measure. For the US, the decay is below five years before the financial crisis. After the financial crisis, as the short rate is hitting the lower bound, the decay becomes much slower to about ten years. For the UK, the decay is within ten years except the spike in late 2010. Except under special circumstances such as right after the crisis, the contribution of the expectation component is limited to the first segment of the swap curve below ten years. After that, one can largely ignore the contribution of the short-rate expectation and focus on the risk premium and convexity contribution, as we have done in the previous section.

[Figure 7 about here.]

For identification, we set the expected rate of change for long-term rates to zero in the limit (μ_∞). As such, μ_0 captures more of the relative component of the expected rate of change between short- and long-maturity swap rates, and thus the swap curve slope changes. To examine whether the extracted rate of change is informative about future swap curve slope changes, Table 4 reports the forecasting correlation estimates between the expected rate of change estimates (μ_0) and changes in the swap curve slope over different horizons (h, in weeks), from one month ($h = 4$ weeks), to one quarter ($h = 13$ weeks), to half a year ($h = 26$) and one year ($h = 52$ weeks). We measure the swap curve slope using the swap rate difference between two year and other maturities from three to ten years. The forecasting correlation estimates are all strongly positive. For the US, the forecasting correlation on changes of the 2-3 swap slope is 27% at one-month horizon, 48% at quarterly horizon, 59% at half-year horizon, and 61% at one-year horizon. The forecasting correlation estimates are similar on other slope measures. For the UK, the forecasting correlation estimates are equally high, from 22% at monthly horizon to 64% at annual horizon on the 2-3

swap slope. The high correlation estimates suggest that the simple common factor structure can effectively separate out an expectation component that is highly predictive of future swap curve slope changes.

3.3. Constructing butterfly positions with local commonality assumptions

Butterfly portfolios constructed with three bonds have been a staple trade in the fixed income market. By hedging away systematic interest-rate movements, a well-constructed butterfly portfolio can become very stable, with only small temporal movements left due to the idiosyncratic movements of the particular bonds.

Bali, Heidari, and Wu (2009) propose to form interest rate portfolios based on estimated DTSMs. These portfolios are constructed to neutralize the common factors so that the remaining fluctuations are driven by the pricing residuals of the model. A three-leg butterfly portfolio can be constructed to neutralize two common factors. If a two-factor structure is well-specified and the pricing residuals are mean-reverting, the constructed butterfly portfolios of interest rates become mean-reverting and one can treat the negative of the pricing residual as the alpha source. The stability and profitability of such portfolios depend crucially on how well the model specification captures the true underlying factor structure. On the other hand, given the centralized nature of DTSMs, such portfolio constructions are in theory agnostic to the particular bond maturity choice. In practice, however, Bali, Heidari, and Wu (2009) show that different maturity combinations can lead to butterfly portfolios with very different behaviors. Some portfolio are very stable and highly mean-reverting, while others are not stable at all.

In this section, based on the new decentralized pricing theory, we propose to construct decentralized butterfly portfolios that focus much less on factor structure assumptions but much more on

the particular choice of maturity combination.

3.3.1. Local commonality of interest rate movements

To motivate our maturity combination choice, we start with the intuitive observation that bonds of *nearby* maturities tend to behave similarly and co-move strongly. The closer the maturities between two bonds, the closer their behaviors are and the stronger their co-movements. To illustrate this behavior, we choose a set of reference maturities and measure the cross-correlations of swap rate changes between these reference maturities and other maturities.

In the US, the most actively traded maturities on swap rates and Treasury bonds center around the maturities of the Treasury bond futures, which are written on bonds at six segments of the maturity spectrum around 3, 5, 7, 10, 15, and 30 years, respectively.[6] We take these maturities as reference maturities and show how their correlations of weeky changes with other maturities decline as the maturity gap increases.

Figure 8 plots the correlation estimates against the maturity gap with the reference maturity. Within each panel, each line plots the correlation of one reference maturity with other maturities. Our data sample includes swap rates at 11 maturities from two to 50 years. The maturity gap measures the gap between the 11 maturities and the reference maturity. For example, for the line with 10-year as the reference maturity, the maturity gaps with 5-, 7-, 10-, 15-, and 20-year swaps are -2, -1, 0, 1, and 2, respectively.

[Figure 8 about here.]

[6]For example, futures with September 2017 delivery include TUU7 (2-year note), FVU7 (5-year note), TYU7 (the old 10-year contract with maturities around 7 years), UXYU7 (the new 10-year ultra, with maturities around 10 year), USU7 (long bond with maturities around 18 years), and WNU7 (ultra bond with maturities over 25 years). We move the first reference maturity to 3-year because the shortest swap-rate maturity we have is at 2-year.

For both US and UK, the correlation estimates decline monotonically as the absolute maturity gap increases. For US, the correlations of weekly changes of swap rates at adjacent maturities (with maturity gaps of ± 1) are higher than 97.8%. The estimates remain high at the next level (with gaps of ± 2) at no less than 95.3%. As the gap increases further, the correlation estimates decline to as low as 67%. For the UK, the corresponding estimates are slightly lower, but remain as high as over 96.6% for swap rates of adjacent maturities, and above 92.4% for the swap rates with gaps of ± 2. The correlation estimates decline to as low as 52.8% as the gap increases further.

We label this behavior as *local commonality*. It is driven less by any model assumptions, but more by the fact that nearby contracts have similar payoffs and should hence behave similarly regardless of any dynamics assumptions.

With this local commonality observation, we propose to build decentralized butterfly portfolios with bonds of nearby maturities. Regardless of the underlying dynamics, since interest rates at nearby maturities show more common movements, butterflies formed with nearby-maturity bonds are likely to become more stable, with only a small proportion of idiosyncratic movements left. By contrast, with maturities far apart, even if the butterflies are supposed to be well hedged under some particular model assumptions, they are unlikely to be stable in practice. It is inherently difficult to construct practically stable butterflies with maturities far apart given the low correlations in their movements.

3.3.2. Constructing decentralized butterflies with nearby maturities

For each butterfly, we normalize the weight at the middle maturity to one, and determine the weights on the two maturities at the wings. We do not make parallel or any other factor dynamics assumptions, but directly estimate the covariance Σ_t of changes in the three swap rate series in the

butterfly portfolio, and choose the butterfly weights to minimize the portfolio variance,

$$\min_{\beta_t} \beta_t^\top \Sigma_t \beta_t, \quad \text{subject to} \quad \beta_t^\top i_2 = 1, \tag{44}$$

where i_2 is an indicator vector with all elements being zero except the second element, as a way of normalizing the weight at the middle maturity to one. The solution is

$$\beta_t = \Sigma_t^{-1} i_2 / (i_2^\top \Sigma_t^{-1} i_2). \tag{45}$$

Instead of inverting the covariance matrix Σ_t, one can also directly regress rate changes at the middle maturity against the two rate changes on the wings. The slope estimates would be for $(-\beta_1, -\beta_3)$.

In the extreme case of perfect parallel movements, the covariance matrix Σ_t becomes singular and the regression can experience the multi-collinearity problem. In this case, we can simply set $\beta_1 = \beta_3 = -0.5$ to construct a symmetric butterfly.

We treat each swap rate as the coupon of a par bond and solve the weight for the par bond by adjusting the interest rate exposure estimates β_t for the duration difference,

$$w_1 = \beta_1 \tau_2 / \tau_1, \quad w_2 = 1, \quad w_3 = \beta_3 \tau_2 / \tau_3. \tag{46}$$

Butterfly portfolios constructed with the above weights have minimum variationa left. Recall the bond return decomposition in (13),

$$\frac{dB_t}{B_t dt} = y_t + \frac{1}{2}\tau^2 \frac{(dy_t)^2}{dt} - \tau \frac{dy_t}{dt}. \tag{47}$$

The carry term is known ex ante, and the convexity term also becomes deterministic under diffusive moves as the term $(dy_t)^2/dt$ converges to the annualized variance rate σ_t^2 of the yield changes. The bond return variation is thus chiefly driven by the variation of the bond yield dy_t. With the butterfly construction, the bond portfolio return becomes,

$$w_t^\top \frac{dB_t}{B_t dt} = w_t^\top y_t + \frac{1}{2} w_t^\top \left(\tau^2 \frac{(dy_t)^2}{dt} \right) - \tau_2 \beta_t^\top \frac{dy_t}{dt}, \tag{48}$$

where the butterfly weight β_t is chosen to minimize the variance of the yield combination $\beta_t^\top dy_t/dt$, given the covariance estimate Σ_t on the vector dy_t/dt.

3.3.3. Maturity gaps and stability of butterfly portfolios

To examine how effective the butterfly construction is in removing systematic interest-rate risk, we estimate the butterfly weights based on full sample regression on weekly changes of the associated swap rates. The regression residual represents the residual risk of the butterfly portfolio $\beta^\top dy$. At each reference maturity, we construct butterflies both with adjacent maturities (with maturity gap of ± 1) and with increasingly larger maturity gaps. We compare how the behaviors of the butterflies vary with the maturity gap of the construction.

Table 5 reports the annualized standard deviation of the weekly changes of the reference swap rates, as well as the variance ratios of weekly changes on the butterfly relative to that on the reference swap rate. The annualized standard deviation measures the magnitude of variation of the reference swap series. The variance ratio highlights how much variation is left through the butterfly construction. We compare the variance ratios of butterflies constructed with increasingly large maturity gaps. The data contain 11 swap series with maturities from two to 50 years, with 10-year maturity as the median maturity. Thus, with 10-year as the reference maturity, the wings of the

butterfly can be constructed with maturity gaps from one to five. With other reference maturities, when one side of the wing maturity reaches the edge of the maturity span, we only increase the maturity gap on the other side. The label shows the absolute maximum maturity gap from one to five.

Panel A of Table 5 reports the variance reduction results on the US swap rates. The reference swap rate series have annualized standard deviations from 0.85 to 0.95 of a percentage point. When we construct butterflies with adjacent maturities, the variance reduction is drastic. The variance estimates of the butterflies are merely 0.5% to 1.4% of the reference swap rate variance. Thus, 98.6-99.5% of the risk is removed via the butterfly construction. As we increase the maturity gap, the variance ratio estimates increase, a sign of deterioration in the effectiveness of variance reduction. With the maturity gap at two, the variance ratios more than double in most maturities, ranging from 1.1% to 2.6%. At the widest maturity gap of five, the variance ratio can reach as high as 7.7%.

Compared to the US swap rates, weekly changes in the UK reference swap rates show lower annualized standard deviation from 0.64 to 0.79 of a percentage point. Furthermore, a larger proportion of the UK interest-rate movements are idiosyncratic as the variance ratios of the UK butterflies are larger. For butterflies constructed with adjacent maturities, the variance ratio estimates range from 1.3% to 3.2%, larger than that for the US case. As the maturity gap increases, the variance ratio estimates increase in most cases and reach as high as 12.4% when the maturity gap reaches the maximum of five.

Another way of examining the stability of the butterflies is to examine its mean-reversion behavior. Table 6 compares the weekly autocorrelation of the weekly changes of the reference swap rates to the weekly autocorrelation of the butterflies. The autocorrelation estimates for the ref-

erence swap rates are close to zero, illustrating the highly persistent mature of these swap rate series.

A well-constructed butterfly should be able to remove the most persistent component of the interest-rate movement and leave a portfolio that is much more reverting. This is indeed the case for the butterflies constructed with adjacent maturities. The weekly autocorrelation estimates range from -22.2% to -51.7% for the US, and from -34.8% to -48.2% for the UK. Thus, butterfly construction with adjacent maturities is equally effective in removing the persistent movements for both currencies. As the maturity gap increases, the effectiveness declines. In particular, when the maturity gap is at the maximum of five, at several reference maturities for both currencies, the weekly autocorrelation estimates for the butterflies become no larger in absolute magnitude than that on the reference swap rate series. Thus, with a large enough maturity gap, the butterfly construction completely loses its effectiveness in removing the persistent component of the swap rate movement.

The statistics in Tables 5 and 6 highlight why butterfly construction is such a common practice in the fixed-income market. By forming a butterfly with nearby maturities, one can effectively remove 97-99% of the variation in the reference swap rate. Furthermore, by hedging away the most persistent component and leaving only the highly mean-reverting variation, the butterfly also exhibits longer term stability, as the variance of the butterfly movements does not increase with the holding horizon as fast as for the individual swap series. Therefore, when there are opposite demands at nearby maturities, broker dealers often use butterfly constructions as a cheaper and more flexible way of managing their interest rate risk. Some investment firms with good broker relations can also share the liquidity provision business with their brokers selectively by timing the strong mean-reversion behavior of the butterflies.

Bali, Heidari, and Wu (2009) propose to construct butterflies to hedge away the two most persistent factors of a DTSM, but they find that the stabilities of the butterflies can vary greatly across different maturity combinations, to the point that they think butterfly construction based on DTSMs is not an effective way of constructing stable interest-rate portfolios. They propose to hedge away more interest-rate factors to achieve universal stability across all maturity combinations. The localized nature of our new theory prompts us to think from a different perspective, and allows us to pay more attention to the maturity choice than the dynamics assumption. Our analysis shows that as long as we choose nearby maturities, butterfly construction can be very effective in removing systematic interest-rate risk, regardless of what the true underlying factor structure is.

Historically, researchers, e.g., Litterman and Scheinkman (1991) and Heidari and Wu (2003), use principal component analysis to determine the number of factors that govern the variation of the interest rate term structure. These findings, however, invariably depend on the number of interest rate series and the maturity span of the data used in the analysis. Fewer factors are needed to explain the variation of interest rates within a narrower maturity span, but as one includes more series with a wider maturity span, the number of factors needed to explain a certain percentage of variation on all series tends to grow. By focusing the local commonality of payoff structures, our new decentralized theory allows us to ignore the global factor structure all together and focus on the strong co-movements of contracts with maturities close to each other.

3.3.4. Out-of-sample investment in decentralized butterflies

The previous section highlights the importance of maturity choice in forming stable butterfly portfolios and shows that decentralized butterflies constructed with nearby maturities are the most stable due to the local commonality of their payoff structures. This section uses an out-of-sample

investment exercise to show how to invest in the stable butterflies to benefit from its strong mean reversion behavior.

Starting on January 3rd, 1996, at each Wednesday t, for each butterfly maturity combination, we use the past one-year history of weekly changes of the swap rates to estimate the annualized variance rate on each series $\widehat{\sigma}_t^2$. We use the variance rate estimate to construct the convexity contribution for each series, $C_t = \frac{1}{2}\tau^2\widehat{\sigma}_t^2$. We also use the one-year history to regress weekly changes of the reference swap rate against weekly changes of the two swap rate series on the wing. The regression slope gives the estimates for $(-\beta_1, -\beta_3)$. From the regression residual series, we obtain an annualized variance rate estimator $\widehat{\sigma}_e^2$ for the butterfly interest-rate portfolio $\beta^\top dy_t$, and use it to construct the conditional variance for the butterfly portfolio bond return as

$$V_t^f = \tau_2^2\widehat{\sigma}_e^2. \tag{49}$$

Given the butterfly weight estimates, we estimate the expected excess return on the portfolio (EER_t^f) by taking expectation on the bond return decomposition equation in equation (13) and deducting it from the financing rate,

$$EER_t^f = w_t^\top(y_t - r_t) + w_t^\top C_t - \tau_2\mu_t^f, \tag{50}$$

where the first term captures the carry contribution over the financing rate, and the second term capture the contribution from convexity. The expected rate of change on each interest rate series $\mu_t = \mathbb{E}_t[dy_t/dt]$ is the most difficult to construct because it is inherently different to predict the highly persistent interest-rate series. Instead of forecasting the expected rate of change on each series separately, we exploit the highly mean-reverting behavior of the butterfly interest-rate portfolio

and directly predict the expected rate of change on the interest-rate portfolio, $\mu_t^f = \mathbb{E}_t\left[\beta_t^\top dy_t/dt\right]$. We take the last day of the regression residual e_t, assume that it will converge to zero in a week, and set the expected rate of change to the negative of this residual annualized to 52 weeks, $\mu_t^f = -52e_t$.

Given the conditional mean and variance estimates, we take long or short positions in the butterfly portfolio based on whether the expected excess return is positive or negative, and we set the investment weight proportional to the conditional mean-variance ratio,

$$n_t = c\,EER_t^f/V_t^f. \tag{51}$$

where c denotes a constant proportional coefficient that we use to scale the weights.

Table 7 reports the summary statistics on the time-varying allocation weights to butterflies constructed with adjacent maturities around each of the six reference maturities, with the last column reporting the average statistics across the fix butterfly portfolios. The statistics include percentile values of the allocation weights at 10, 25, 50, 75, and 90 percentiles. The median allocation weights are close to zero and the percentile values show a reasonably symmetric distribution for the allocation weights. For both currencies, the variations of the allocation weights are larger at shorter maturities than at longer maturities. Since we apply the same proportionality coefficient c, the different degree of variation across maturities reflects the duration difference of the reference maturities. Given the same regression residual, the allocation weight is smaller for longer-duration butterflies. Since the butterflies have very little risk left, investors often take highly levered positions in their butterfly investments. While the choice of the constant scaling coefficient c is immaterial to our analysis, we choose a scaling level so that the variations of the weights are within the usual confines of the leverage levels employed by investors.

The last row of each panel reports the average cross-correlation estimates of each allocation

weight series with the other allocation weights. The cross-correlations between the allocation weights to different butterfly portfolios are close to zero, suggesting that the variations of the expected excess returns from these localized butterfly portfolios are largely independent of each other. This behavior forms a sharp contrast with the extremely high cross-correlation between the market price of risk estimates extracted from different par bonds in Section 3.1. While the expected excess returns on different par bonds tend to move strongly together, the expected excess returns on our localized butterfly portfolios are largely independent of one another.

Table 8 reports the summary statistics of the weekly excess returns from investing in the localized butterflies constructed with adjacent maturities. The investment performances are similar for butterflies around different reference maturities and for both currencies. The annualized mean excess returns range from 9% to 23%. The annualized standard deviations range from 7% to 16%. The annualized information ratios range from 1.03 to 1.47. The excess returns show large excess kurtosis, and moderately positive skewness. Just as the allocation weights across different butterflies are largely independent of each other, so are the realized excess returns from different butterfly portfolios. The near-independent nature of the decentralized butterflies offers strong diversification benefits when one simultaneously invests in all the six butterfly portfolios. The last column reports the excess return statistics from an equal weighted portfolio of the six butterflies. While the mean excess return retains the average magnitude of 12-13%, the standard deviation of the aggregate portfolio becomes much smaller than that of the individual butterflies at merely 5-6%. The annualized information nearly doubles that of the individual butterfly investment and reaches 2.32 for the UK portfolio and 2.66 for the US portfolio.

For comparison, we also repeat the investment exercise on butterflies constructed with larger maturity gaps. Table 9 reports the summary statistics of the excess returns from investing in butterflies constructed with maturity gaps of two. Across all reference maturities and both currencies,

the annualized information ratios of the investments become significantly lower. For most butterflies, the performance reduction comes from both a smaller mean excess return and a larger standard deviation. For the aggregate butterfly portfolio, the annualized information ratio declines by 40% from 2.66 to 1.66 in the US and by 25% from 2.32 to 1.74 in the UK. Furthermore, with the larger maturity gap, the excess returns show larger excess kurtosis but smaller or even negative skewness. Thus, as the maturity gap increases, the investment not only shows worse average performance, but the performance also becomes less stable, and is more likely to experience large negative realizations.

When the maturity gap increases further, butterflies around different reference maturities become intertwined on the wings and thus lose their decentralized nature. Not only do the butterflies become less stable and their investment performances deteriorate, but the investment returns also become more correlated with one another, losing further on the benefits of diversification. When we increase the maturity gap to five in the butterfly construction, the butterfly portfolios at different reference maturities become completely intertwined, and the investment returns from the six butterflies show an average correlation of 0.34 in the US and 0.37 in the UK, much higher than the return correlations for the butterflies constructed with adjacent maturities. The annualized information ratio for the aggregate butterfly portfolio is reduced to 1.09 for the US portfolio and even lower at 0.58 for UK portfolio, largely losing the allure of the staple trade.

4. Concluding Remarks

In this paper, we propose a new modeling framework that is particularly suited for analyzing returns on a bond or bond portfolio. The framework does not try to model the full dynamics of an instantaneous short rate, but focuses squarely on the behavior of the bond yield in question. It does

not even ask for the full dynamics specification of this bond yield, but only needs estimates of its conditional expected rate of change, risk premium, and volatility.

We explore applications for the new theory from several angles. We show that we can predict bond excess returns on long-term bonds, without running predictive regressions, even by assuming no prediction on long-term rates. We also show how to perform comparative analysis on the yield curve via common factor structures on the near-term dynamics. Finally, the decentralized nature of our theory prompts us to pay more attention to maturity combination than dynamics assumptions in constructing the popular butterfly bond portfolios. We show that decentralized butterfly portfolios constructed with nearby maturities show much more stable behavior than those constructed with maturities far apart.

For future research, separating risk premium from expectation via common factor structure assumptions and/or new information sources remains a challenging, but fruitful endeavor. The new decentralized theory allows one to consider the separation based on localized information, potentially offering a more convenient way of exploring domain expertise.

References

Adrian, T., R. K. Crump, and E. Moench. 2013. Pricing the Term Structure with Linear Regressions. *Journal of Financial Economics* 110(1):110–138.

Aït-Sahalia, Y. 1996. Nonparametric Pricing of Interest Rate Derivatives. *Econometrica* 64(3):527–560.

Aït-Sahalia, Y, and R. Kimmel. 2010. Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions. *Journal of Financial Economics* 98(1):113–144.

Bali, T., M. Heidari, and L. Wu. 2009. Predictability of Interest Rates and Interest-Rate Portfolios. *Journal of Business and Economic Statistics* 27(4):517–527.

Bekaert, G., R. J. Hodrick, and D. A. Marshall. 1997. On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates. *Journal of Financial Economics* 44(3):309–348.

Black, F. 1995. Interest Rates as Options. *Journal of Finance* 50(5):1371–1376.

Brace, A., D. Gatarek, and M. Musiela. 1997. The Market Model of Interest Rate Dynamics. *Mathematical Finance* 7(2):127–255.

Campbell, J. Y. 1995. Some Lessons from the Yield Curve. *Journal of Economic Perspectives* 9(3):129–152.

Campbell, J. Y, and R. Shiller. 1991. Yield Spreads and Interest Rate Movements: A Bird'S Eye View. *Review of Economic Studies* 58(3):494–514.

Cheng, P, and O. Scaillet. 2007. Linear-Quadratic Jump-Diffusion Modelling with Application to Stochastic Volatility. *Mathematical Finance* 17:575–598.

Christensen, J. H. E, and G. D. Rudebusch. 2015. Estimating Shadow-Rate Term Structure Models with Near-Zero Yields. *Journal of Financial Econometrics* 13(2):226–259.

Cieslak, A, and P. Povala. 2015. Expected Returns in Treasury Bonds. *Review of Financial Studies* 28(10):2859–2901.

Cochrane, J. H, and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review* 95(1):138–160.

Collin-Dufresne, P., R. S. Goldstein, and C. S. Jones. 2008. Identification of Maximal Affine Term Structure Models. *Journal of Finance* 63(2):743–795.

Cooper, I, and R. Priestley. 2009. Time-Varying Risk Premiums and the Output Gap. *Review of Financial Studies* 22(7):2801–2833.

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

Dai, Q, and K. Singleton. 2002. Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure. *Journal of Financial Economics* 63:415–441.

Duarte, J., F. A. Longstaff, and F. Yu. 2007. Risk and Return in Fixed Income Arbitrage: Nickels in Front of a Streamroller. *Review of Financial Studies* 20(3):769–811.

Duffee, G. R. 2002. Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance* 57(1):405–443.

Duffie, D., D. Filipović, and W. Schachermayer. 2003. Affine Processes and Applications in Finance. *Annals of Applied Probability* 13(3):984–1053.

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

Duffie, D., J. Pan, and K. Singleton. 2000. Transform Analysis and Asset Pricing for Affine Jump Diffusions. *Econometrica* 68(6):1343–1376.

Dybvig, P. H., J. E. Ingersoll, and S. A. Ross. 1996. Long Forward And Zero-Coupon Rates Can Never Fall. *Journal of Business* 69(1):1–25.

Fama, E. F, and K. R. French. 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics* 25(1):23–49.

Gargano, A., D. Pettenuzzo, and A. Timmermann. 2017. Bond Return Predictability: Economic Value and Links to the Macroeconomy. *management Science* 65(2):508–540.

Giglio, S, and B. T. Kelly. 2018. Excess Volatility: Beyond Discount Rates. *Quarterly Journal of Economics* 133(1):71–127.

Goldstein, R. 2000. The Term Structure of Interest Rates as a Random Field. *Review of Financial Studies* 13(2):365–384.

Hamilton, J. D, and J. C. Wu. 2012. Identification and Estimation of Gaussian Affine Term Structure Models. *Journal of Econometrics* 168(2):315–331.

Heath, D., R. Jarrow, and A. Morton. 1992. Bond Pricing and the Term Structure of Interest Rates: A New Technology for Contingent Claims Valuation. *Econometrica* 60(1):77–105.

Heidari, M, and L. Wu. 2003. Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?. *Journal of Fixed Income* 13(1):75–86.

Ho, T. S. Y, and S. B. Lee. 1986. Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance* 41(5):1011–1030.

Jamshidian, F. 1997. LIBOR and Swap Market Models and Measures. *Finance and Stochastics* 1(4):293–330.

Joslin, S., K. J. Singleton, and H. Zhu. 2011. A New Perspective on Gaussian Dynamic Term Structure Models. *Review of Financial Studies* 24(3):926–970.

Kalman, R. E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME–Journal of Basic Engineering* 82(Series D):35–45.

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

Krippner, L. 2012. Measuring the Stance of Monetary Policy in Zero Lower Bound Environments. *Economics Letters* 118(1):135–138.

Krippner, L. 2013. A tractable Framework for Zero Lower Bound Gaussian Term Structure Models. Discussion paper Reserve Bank of New Zealand.

Leippold, M, and L. Wu. 2002. Asset Pricing under the Quadratic Class. *Journal of Financial and Quantitative Analysis* 37(2):271–295.

Leippold, M, and L. Wu. 2003. Design and Estimation of Quadratic Term Structure Models. *European Finance Review* 7(1):47–73.

Litterman, R, and J. Scheinkman. 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income* 1(1):54–61.

Longstaff, F. A. 2000. The Term Structure of Very Short-Term Rates: New Evidence for the Expectation Hypothesis. *Journal of Financial Economics* 58(3):397–415.

Ludvigson, S. C, and S. Ng. 2009. Macro Factors in Bond Risk Premia. *Review of Financial Studies* 22(12):5027–5067.

Merton, R. C. 1973. Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science* 4(1):141–183.

Miltersen, K. R., K. Sandmann, and D. Sondermann. 1997. Closed-form Solutions for Term Structure Derivatives with Lognormal Interest Rates. *Journal of Finance* 52(1):409–430.

Santa-Clara, P, and D. Sornette. 2001. The Dynamics of Forward Interest Rate Curve with Stochastic String Shocks. *Review of Financial Studies* 14(1):149–185.

Wan, E. A, and R. van der Merwe. 2001. The Unscented Kalman Filter. in *Kalman Filtering and Neural Networks*, ed. by S. Haykin. Wiley & Sons Publishing New York chap. 7, pp. 221–280.

Table 1
Summary statistics of swap rates

	US					UK				
Maturity	Mean	Stdev	Skew	Kurt	Auto	Mean	Stdev	Skew	Kurt	Auto
Panel A. Statistics on swap rate levels										
2	3.24	2.26	0.24	-1.46	0.998	3.95	2.44	-0.10	-1.50	0.998
3	3.49	2.18	0.18	-1.41	0.997	4.10	2.38	-0.11	-1.41	0.997
4	3.70	2.10	0.13	-1.37	0.997	4.21	2.32	-0.11	-1.31	0.997
5	3.89	2.02	0.10	-1.34	0.997	4.30	2.26	-0.10	-1.21	0.997
7	4.17	1.90	0.05	-1.28	0.996	4.44	2.15	-0.05	-1.02	0.997
10	4.45	1.79	-0.00	-1.22	0.996	4.57	2.03	0.04	-0.76	0.996
15	4.72	1.72	-0.07	-1.18	0.996	4.28	1.57	-0.33	-0.75	0.995
20	4.83	1.69	-0.11	-1.17	0.996	4.25	1.46	-0.39	-0.60	0.995
30	4.89	1.66	-0.12	-1.14	0.996	3.99	1.28	-0.64	-0.55	0.996
40	3.76	1.14	0.18	-1.29	0.994	3.86	1.23	-0.69	-0.40	0.995
50	3.74	1.15	0.19	-1.29	0.994	3.46	1.12	-0.77	-0.56	0.995
Panel B. Statistics on swap rate weekly changes										
2	-0.26	0.82	0.19	3.44	0.011	-0.34	0.75	0.07	3.30	-0.009
3	-0.26	0.89	0.25	2.42	-0.003	-0.34	0.79	0.30	2.76	-0.036
4	-0.26	0.92	0.26	1.86	-0.011	-0.34	0.79	0.29	2.12	-0.058
5	-0.26	0.94	0.24	1.72	-0.020	-0.34	0.78	0.21	1.68	-0.061
7	-0.25	0.95	0.24	1.72	-0.028	-0.34	0.76	0.21	1.44	-0.080
10	-0.25	0.94	0.21	1.98	-0.035	-0.34	0.74	0.16	1.27	-0.083
15	-0.25	0.90	0.24	1.81	-0.051	-0.32	0.70	0.10	1.45	-0.094
20	-0.25	0.88	0.20	1.61	-0.050	-0.29	0.67	0.07	1.75	-0.102
30	-0.25	0.85	0.20	1.83	-0.033	-0.20	0.64	-0.15	2.49	-0.105
40	-0.21	0.87	0.09	2.71	-0.048	-0.22	0.64	-0.17	2.78	-0.100
50	-0.21	0.88	0.04	2.78	-0.053	-0.24	0.60	-0.41	3.61	-0.077

Entries report the summary statistics of the US and UK swap rates. Data are weekly from January 4, 1995 to December 27th, 2017, for 1200 weeks. US 40- and 50-year swap rates start at a later date with 685 weekly observations each. Longer-maturity UK swap rates also start at later dates, with 1081, 1074, 989, 956, 751 weekly observations at 15-, 20-, 30-, 40-, and 50-year maturity, respectively. The statistics include the sample average ("Mean"), the standard deviation ("Stdev"), the skewness ("Skew"), the excess kurtosis ("Kurt"), and weekly autocorrelation ("Auto"). Panel A reports the statistics on the interest rate levels and Panel B reports the statistics on weekly differences in the swap rates, with the mean and standard deviation of the weekly changes being annualized.

Table 2
Forecasting long bond excess returns with no rate predictability

Maturity	10	15	20	30	40	50
Panel A. Bond return horizon: 6-month						
US	0.31	0.28	0.26	0.23	0.28	0.26
UK	0.20	0.19	0.17	0.21	0.19	0.19
Panel B. Bond return horizon: One year						
US	0.37	0.36	0.33	0.30	0.35	0.33
UK	0.35	0.34	0.31	0.33	0.31	0.33

Entries report the forecasting correlation between the bond risk premium extracted from each swap rate and the future excess return of the corresponding par bond over the next six months (panel A) and one year (panel B). Each column denotes one swap series.

Table 3
Pricing performance of the common factor structure

	US				UK			
Maturity	Mean	RMSE	Auto	EV	Mean	RMSE	Auto	EV
2	-1.79	9.02	0.848	0.966	-0.55	13.58	0.923	0.914
3	-2.76	10.12	0.896	0.981	-2.26	9.19	0.870	0.983
4	-5.16	10.60	0.893	0.989	-3.95	7.98	0.847	0.993
5	-9.42	12.66	0.884	0.994	-5.38	10.56	0.932	0.992
7	-3.46	8.04	0.884	0.997	-0.73	8.19	0.941	0.996
10	3.43	10.29	0.949	0.996	5.44	9.93	0.949	0.997
15	12.17	16.34	0.972	0.996	10.95	15.07	0.964	0.996
20	8.03	11.59	0.954	0.998	8.27	10.91	0.927	0.998
30	5.14	7.96	0.878	0.999	0.99	6.51	0.897	0.999
40	-4.62	7.64	0.910	0.999	-6.16	10.57	0.934	0.998
50	-17.36	21.89	0.973	0.995	-12.43	17.55	0.962	0.996

Entries report the summary statistics of the pricing errors of the common factor structure on the US and UK swap rates, including the sample average ("Mean") in basis points, root mean squared pricing error ("RMSE") in basis points, the weekly autocorrelation ("Auto"), and the model's explained variation (EV), defined as one minus the variance ratio of the pricing error to the original data series, and

Table 4
Predicting changes in yield curve slope with the extracted rate of change

	Swap-curve slope				
Horizon, *h*	2-3	2-4	2-5	2-7	2-10
Panel A. US					
4	0.27	0.27	0.26	0.25	0.24
13	0.48	0.45	0.44	0.41	0.39
26	0.59	0.56	0.54	0.51	0.48
52	0.61	0.58	0.56	0.53	0.50
Panel B. UK					
4	0.22	0.19	0.16	0.12	0.09
13	0.40	0.35	0.29	0.22	0.15
26	0.59	0.52	0.45	0.35	0.26
52	0.64	0.60	0.55	0.48	0.39

Entries report the forecasting correlation between the extracted rate of change μ_0 and future changes in the swap curve slope over different horizons h, measured in number of weeks. The swap curve slopes are measured as swap rate difference between the two-year and other maturities from three to ten years.

Table 5
Effectiveness of variance reduction via butterfly construction

Reference	Swap	Variance ratio of butterflies with maturity gap:				
Maturity	Stdev	1	2	3	4	5
Panel A. US						
3	0.89	0.008	0.011	0.014	0.020	0.023
5	0.94	0.008	0.016	0.044	0.053	0.065
7	0.95	0.006	0.016	0.032	0.068	0.076
10	0.94	0.014	0.026	0.043	0.052	0.077
15	0.90	0.005	0.015	0.019	0.036	0.043
30	0.85	0.007	0.020	0.030	0.033	0.035
Panel B. UK						
3	0.79	0.019	0.022	0.027	0.032	0.036
5	0.78	0.018	0.038	0.086	0.104	0.118
7	0.76	0.013	0.026	0.066	0.124	0.124
10	0.74	0.016	0.037	0.070	0.099	0.116
15	0.70	0.032	0.038	0.062	0.074	0.091
30	0.64	0.026	0.023	0.025	0.027	0.029

Entries report the annualized standard deviation of weekly changes on each reference swap rate series (Stdev), as well as the variance ratio of the butterflies constructed around the reference maturity relative to the reference swap rate itself. For each reference maturity, butterflies are constructed with increasing maturity gaps, with the maturity gap of one indicating construction with adjacent maturities. When one side of the wing maturity reaches the edge of the maturity choice, the maturity gap increase reflects the maturity choice of the other wing.

Table 6
Mean reversion in swap rates and butterfly portfolios

Reference	Changes in	Butterflies with maturity gap:				
Maturity	Swap	1	2	3	4	5
Panel A. US						
3	-0.003	-0.394	-0.334	-0.227	-0.177	-0.156
5	-0.020	-0.517	-0.227	-0.076	-0.037	-0.018
7	-0.028	-0.435	-0.249	-0.100	-0.042	-0.012
10	-0.035	-0.401	-0.317	-0.248	-0.115	-0.165
15	-0.051	-0.438	-0.324	-0.304	-0.343	-0.320
30	-0.033	-0.222	-0.152	-0.273	-0.323	-0.358
Panel B. UK						
3	-0.036	-0.464	-0.379	-0.308	-0.226	-0.174
5	-0.061	-0.455	-0.298	-0.133	-0.093	-0.065
7	-0.080	-0.426	-0.263	-0.211	-0.068	-0.051
10	-0.083	-0.348	-0.212	-0.138	-0.096	-0.016
15	-0.094	-0.482	-0.194	-0.122	-0.092	-0.102
30	-0.104	-0.358	-0.299	-0.309	-0.290	-0.283

Entries report the weekly autocorrelation estimates of weekly changes on each reference swap rate series, as well as on butterflies constructed around the reference maturity. For each reference maturity, butterflies are constructed with increasing maturity gaps, with the maturity gap of one indicating construction with adjacent maturities. When one side of the wing maturity reaches the edge of the maturity choice, the maturity gap increase reflects the maturity choice of the other wing.

Table 7
Summary statistics of the time-varying allocation weights to the butterfly portfolios

Maturity	3	5	7	10	15	30	Average
Panel A. US							
p10	-39.65	-18.60	-20.31	-13.72	-15.27	-10.57	-19.69
p25	-15.23	-7.58	-7.24	-4.41	-5.03	-3.79	-7.21
p50	-0.55	-0.15	-0.13	0.09	0.13	0.13	-0.08
p75	14.43	6.58	7.09	4.47	5.48	4.64	7.11
p90	42.35	17.70	20.17	12.50	16.61	12.37	20.28
Corr	-0.03	-0.05	-0.04	-0.03	0.01	0.01	-0.02
Panel B. UK							
p10	-28.10	-16.33	-11.83	-8.37	-6.37	-5.16	-12.69
p25	-9.89	-7.36	-5.06	-3.79	-2.30	-2.07	-5.08
p50	-0.66	-0.27	-0.00	-0.09	-0.04	0.11	-0.16
p75	9.51	6.22	5.00	3.80	2.46	2.23	4.87
p90	26.76	15.25	12.90	8.47	6.78	4.92	12.51
Corr	-0.01	-0.03	-0.16	-0.13	-0.05	-0.02	-0.07

Entries report the summary statistics of the allocation weights to butterflies constructed with adjacent maturities around each of the six reference maturity. The statistics include the values of the allocation weights at 10, 25, 50, 75, and 90 percentiles. The last row of each panel reports the average cross-correlation of the allocation weights of each butterfly with other butterflies. The last column reports the average statistics across the six butterflies.

Table 8
Investment excess returns on localized butterfly portfolios with adjacent maturities

Maturity	3	5	7	10	15	30	Aggregate
Panel A. US							
Mean	0.13	0.12	0.09	0.09	0.14	0.20	0.12
Stdev	0.10	0.09	0.07	0.09	0.08	0.16	0.05
IR	1.26	1.44	1.17	1.03	1.61	1.24	2.66
Skewness	2.33	2.87	4.25	1.37	3.65	2.06	0.99
Kurtosis	32.07	27.25	42.38	56.78	33.86	36.11	24.32
Corr	-0.01	0.08	0.11	0.10	0.06	0.09	0.07
Panel B. UK							
Mean	0.14	0.12	0.10	0.10	0.12	0.23	0.13
Stdev	0.11	0.09	0.07	0.08	0.10	0.16	0.06
IR	1.24	1.24	1.42	1.35	1.25	1.47	2.32
Skewness	2.30	6.22	2.42	6.77	4.37	7.99	4.50
Kurtosis	63.88	73.29	54.24	103.34	40.03	125.42	45.49
Corr	0.07	0.19	0.28	0.25	0.25	0.13	0.20

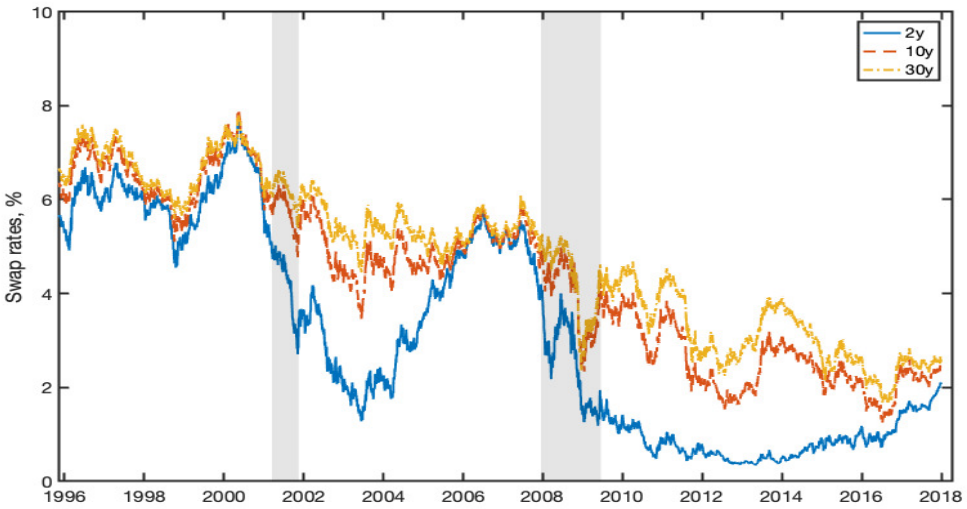
Entries report the summary statistics of excess returns from investments on butterfly portfolios constructed with adjacent maturities. The statistics include the annualized mean excess return (Mean), the annualized standard deviation (Stdev), the annualized information ratio (IR), the skewness (Skewness), the excess kurtosis (Kurtosis), and the average correlation of each excess return series with the other series (Corr). The last column reports the statistics of an aggregate portfolio constructed with equal weighting on the six butterfly portfolios, and the correlation value in the last column represents the grand average of the cross-correlations among the six butterfly investment returns.

Table 9
Investment excess returns on butterfly portfolios with maturity gaps of two

Maturity	3	5	7	10	15	30	Aggregate
Panel A. US							
Mean	0.11	0.10	0.06	0.06	0.10	0.21	0.10
Stdev	0.12	0.10	0.09	0.10	0.11	0.23	0.06
IR	0.97	0.97	0.67	0.57	0.85	0.94	1.60
Skewness	2.05	1.31	-1.12	-4.37	0.24	3.68	-3.29
Kurtosis	24.82	28.98	35.31	93.36	45.12	52.83	73.57
Corr	-0.03	0.09	0.26	0.27	0.21	0.06	0.14
Panel B. UK							
Mean	0.12	0.08	0.06	0.04	0.06	0.16	0.08
Stdev	0.11	0.08	0.07	0.07	0.08	0.15	0.05
IR	1.10	1.04	0.84	0.66	0.81	1.04	1.73
Skewness	1.93	3.58	2.63	-3.44	3.47	2.20	1.84
Kurtosis	72.91	47.21	25.24	79.32	78.32	66.86	34.37
Corr	0.13	0.08	0.23	0.13	0.22	0.10	0.15

Entries report the summary statistics of excess returns from investments on butterfly portfolios constructed with maturity gaps of two. The statistics include the annualized mean excess return (Mean), the annualized standard deviation (Stdev), the annualized information ratio (IR), the skewness (Skewness), the excess kurtosis (Kurtosis), and the average correlation of each excess return series with the other series (Corr). The last column reports the statistics of an aggregate portfolio constructed with equal weighting on the six butterfly portfolios, and the correlation value in the last column represents the grand average of the cross-correlations among the six butterfly investment returns.

Panel A. US



Panel B. UK

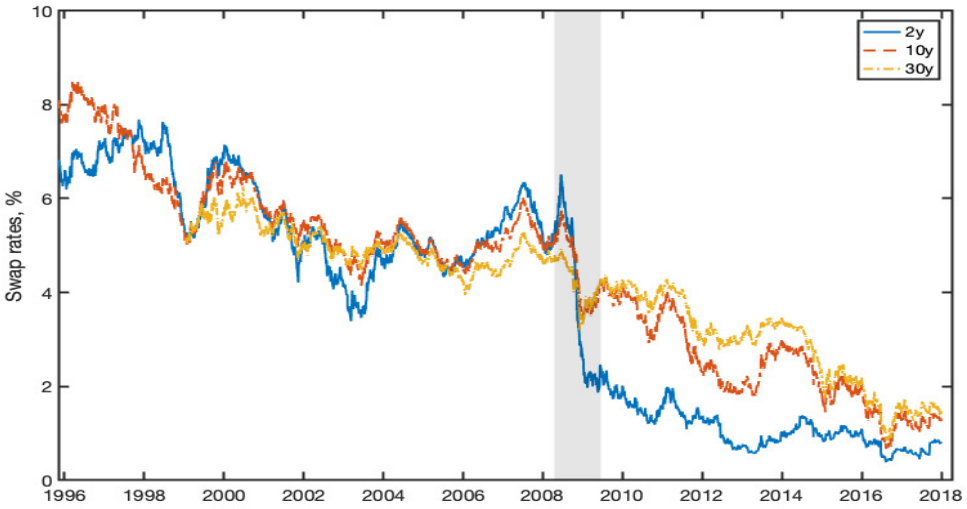
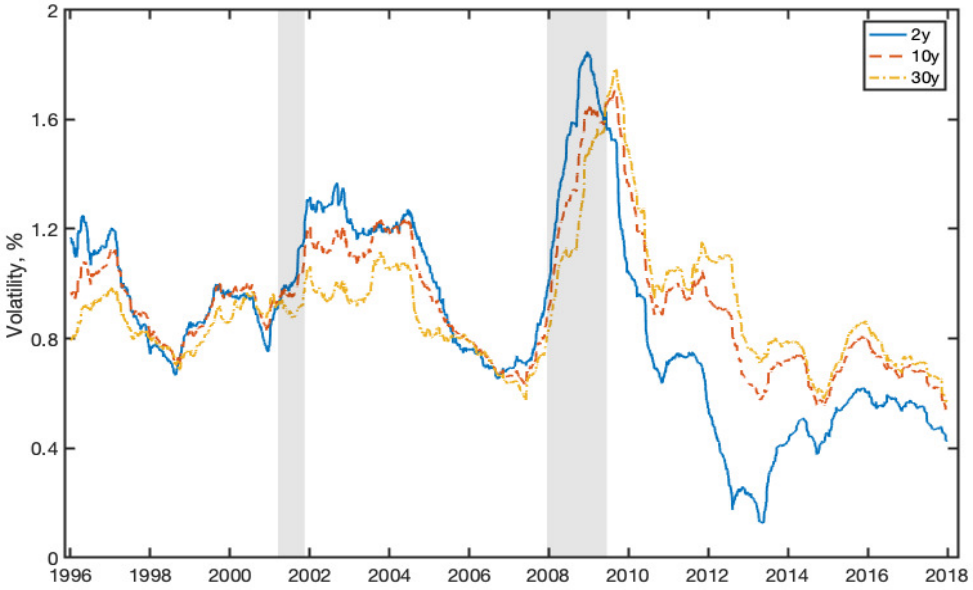


Figure 1

The time-series variation of US and UK swap rates

Each panel plots the time series of swap rates at three selected maturities: 2-year (solid line), 10-year (dashed line), and 30-year (dash-dotted line), overlaid with the recession band of the corresponding economy. Panel A represents the US swap rates and Panel B the UK swap rates.

Panel A. US



Panel B. UK

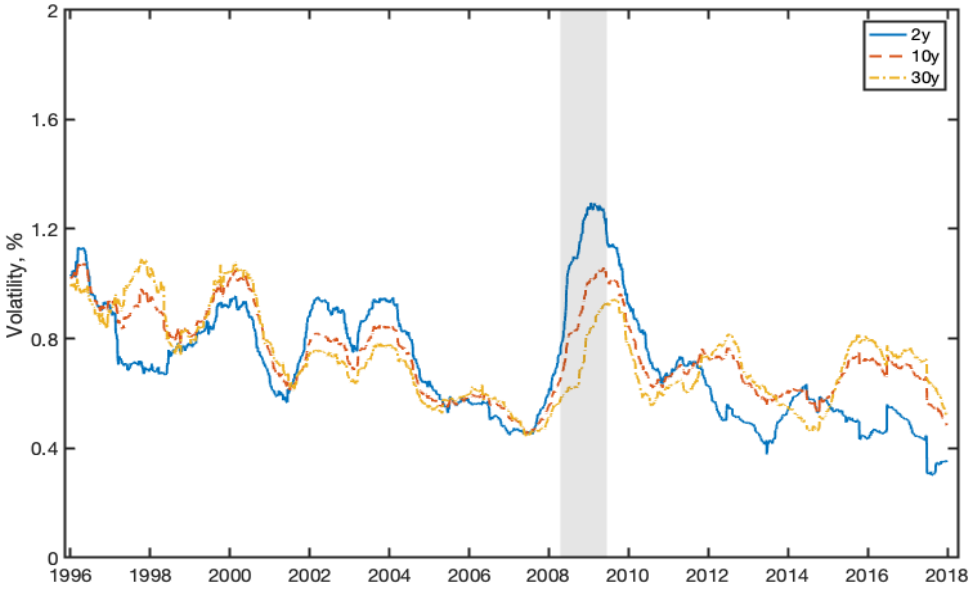
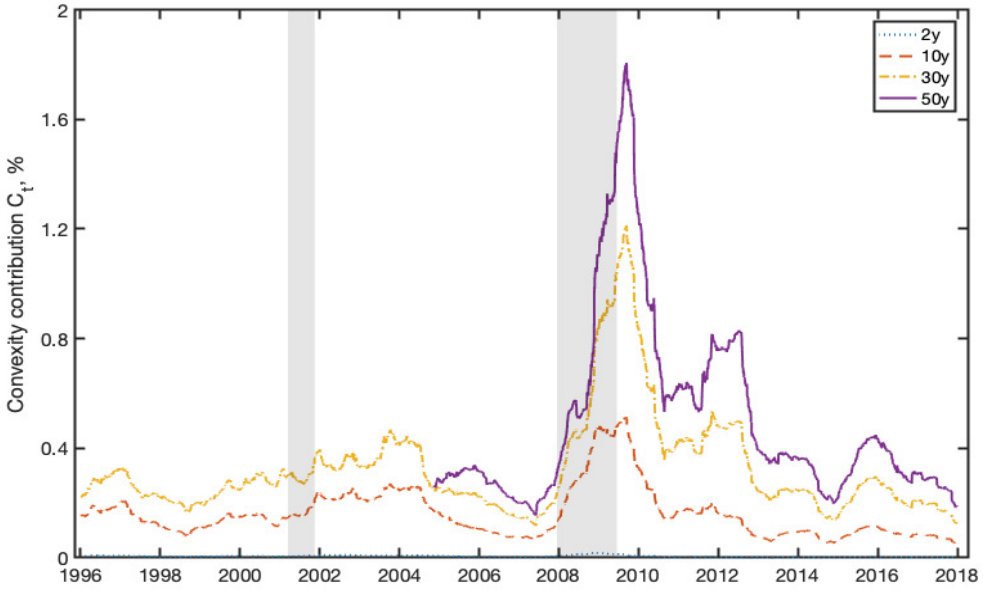


Figure 2
Volatility estimators on US and UK swap rates
Each panel plots the time series of the one-year rolling volatility estimators on the daily changes of the swap rate series at selected maturities: 2-year (solid line), 10-year (dashed line), and 30-year (dash-dotted line), overlaid with the recession band of the corresponding economy. Panel A represent the US swap rates and Panel B the UK swap rates.

Panel A. US



Panel B. UK

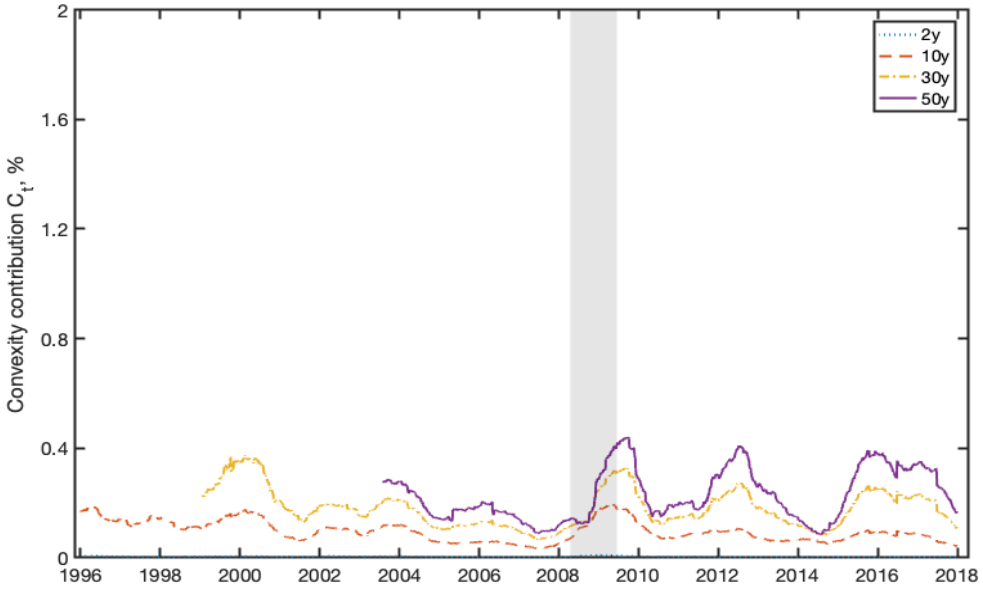
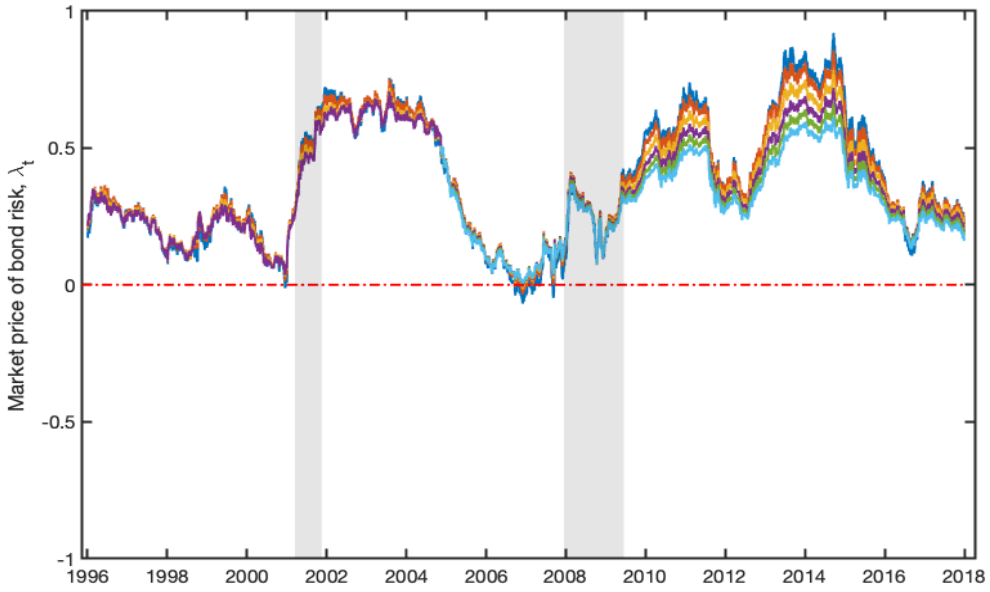


Figure 3
Convexity contribution at selected maturities
Lines denote the time series of the estimated convexity contribution to swap rates at selected maturities: 2-year (dotted line), 10-year (dashed line), 30-year (dash-dotted line), and 50-year (solid line), overlaid with the recession band of the corresponding economy, US in panel A and UK in panel B.

Panel A. US



Panel B. UK

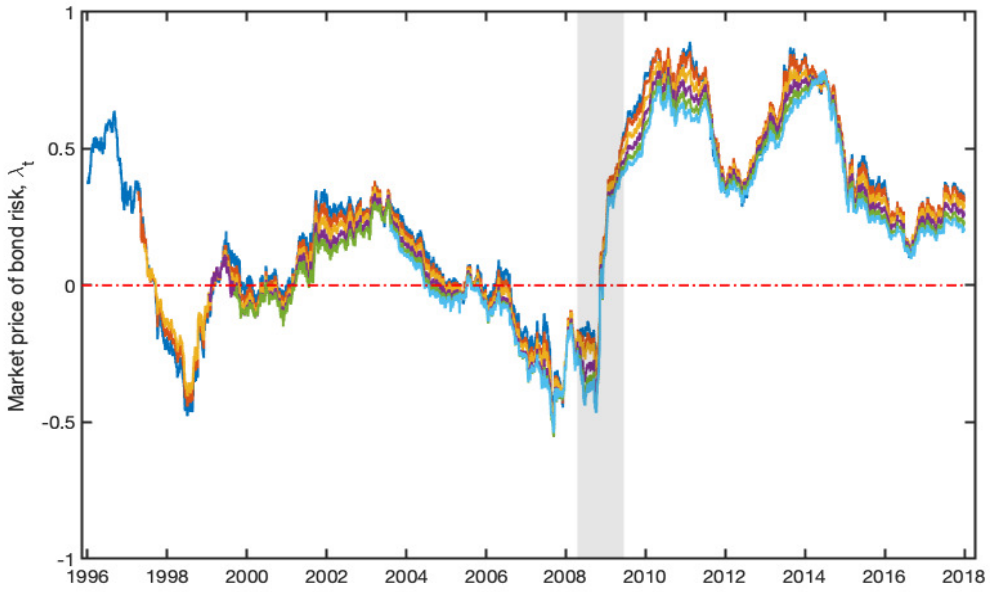
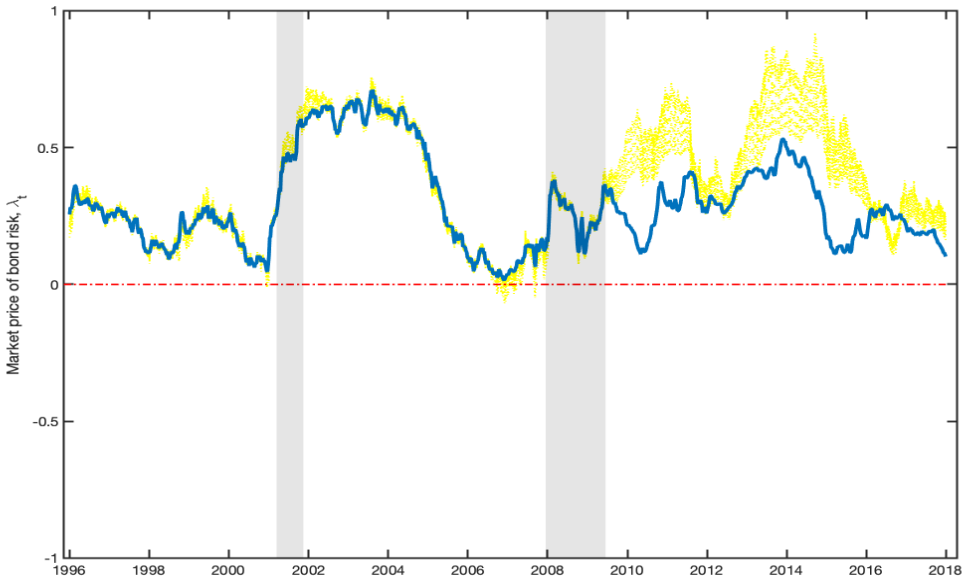


Figure 4

Market price of bond risk extracted from no-predictability assumption

Lines denote the time series of the market price of bond risk extracted from swap rates with maturities 10 years and longer, with the assumption of no rate predictability, overlaid with the recession bands. Panel A is for the US and Panel B is for the UK.

Panel A. US



Panel B. UK

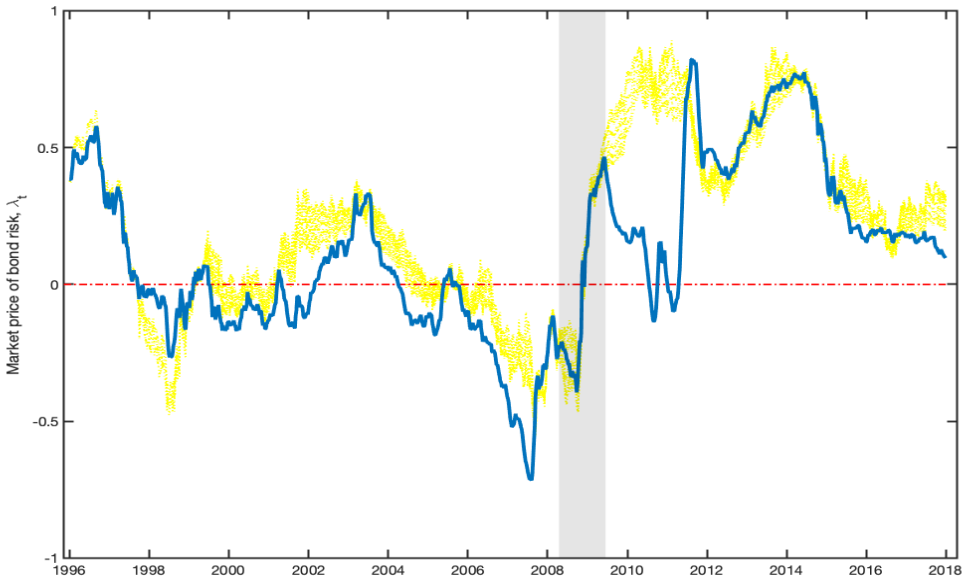
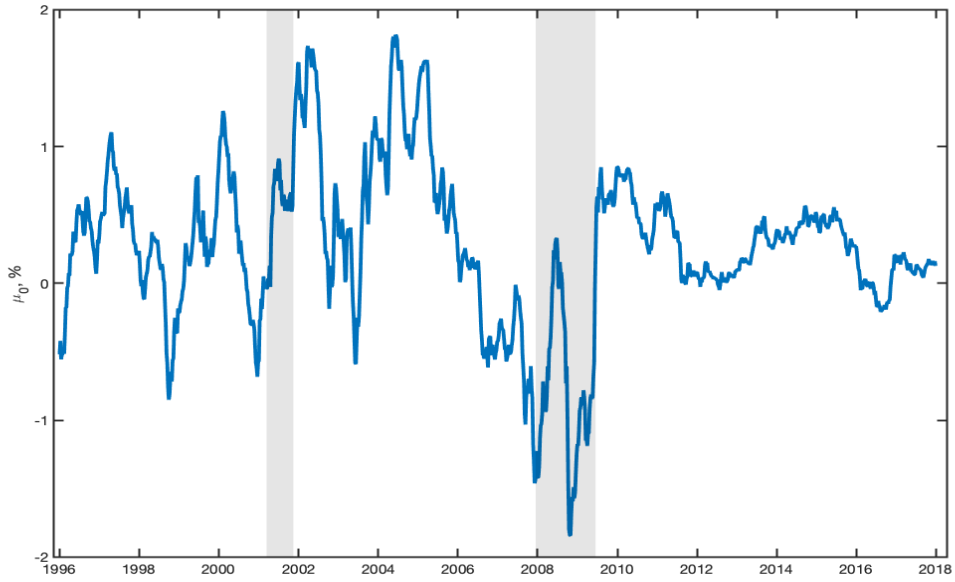


Figure 5

The market price of bond risk extracted from a common factor structure

The solid line in each panel plots the time series of the market price of bond risk, extracted from the common factor structure on the swap rate curve, overlaid with the recession bands. The dashed lines are estimates from the previous subsection based on no-predictability assumption on long-term rates. Panel A is for the US and Panel B is for the UK.

Panel A. US



Panel B. UK

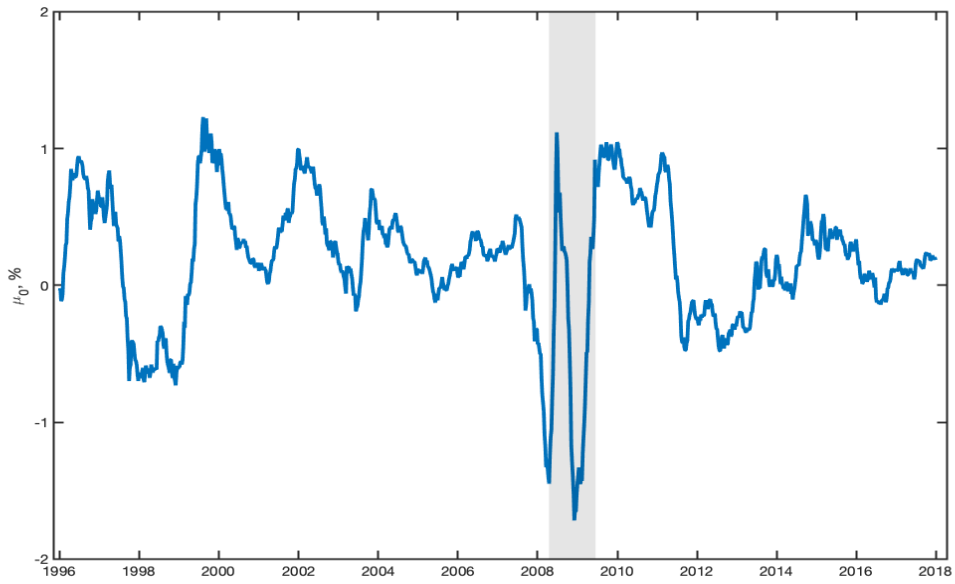
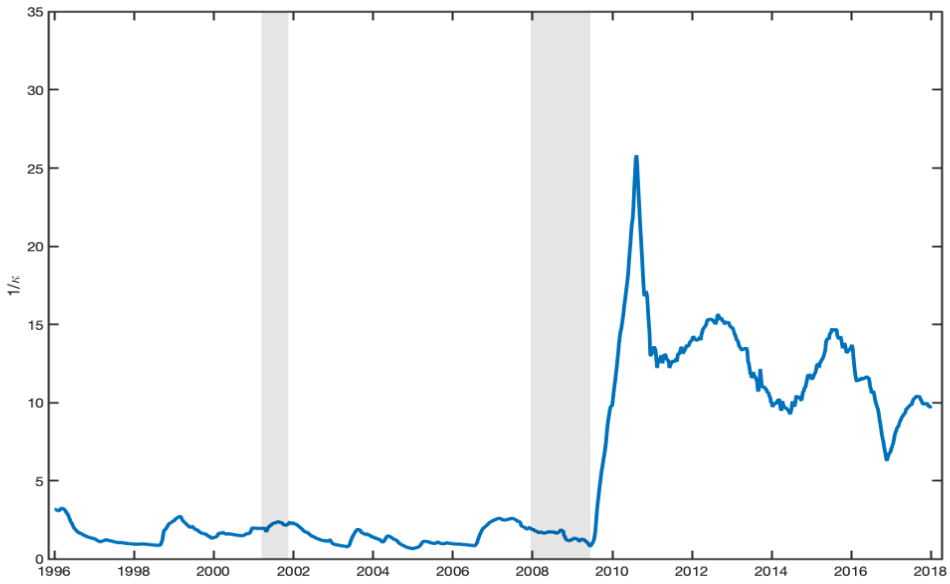


Figure 6
The expected rate of change μ_0
Lines plot the time series of the expected rate of change on short rates μ_0 extracted from the common factor structure on the swap rate curve. The plots are overlaid with the recession bands. Panel A is for the US and Panel B is for the UK.

Panel A. US



Panel B. UK

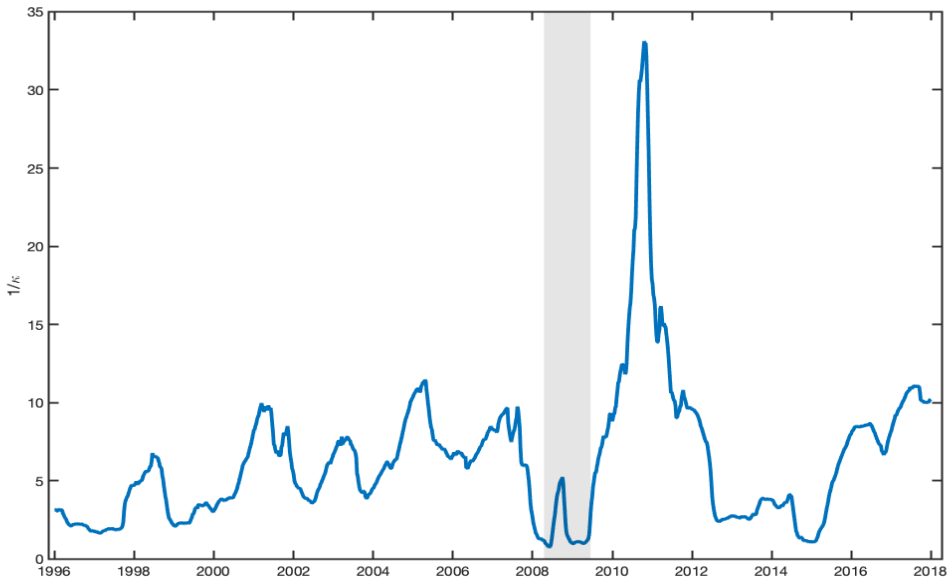


Figure 7
The reciprocal of the decay speed, $1/\kappa$
Lines plot the time series of the reciprocal of the decay speed, $1/\kappa$, extracted fro the common factor structure on the swap rate curve. The plots are overlaid with the recession bands. Panel A is for the US and Panel B is for the UK.

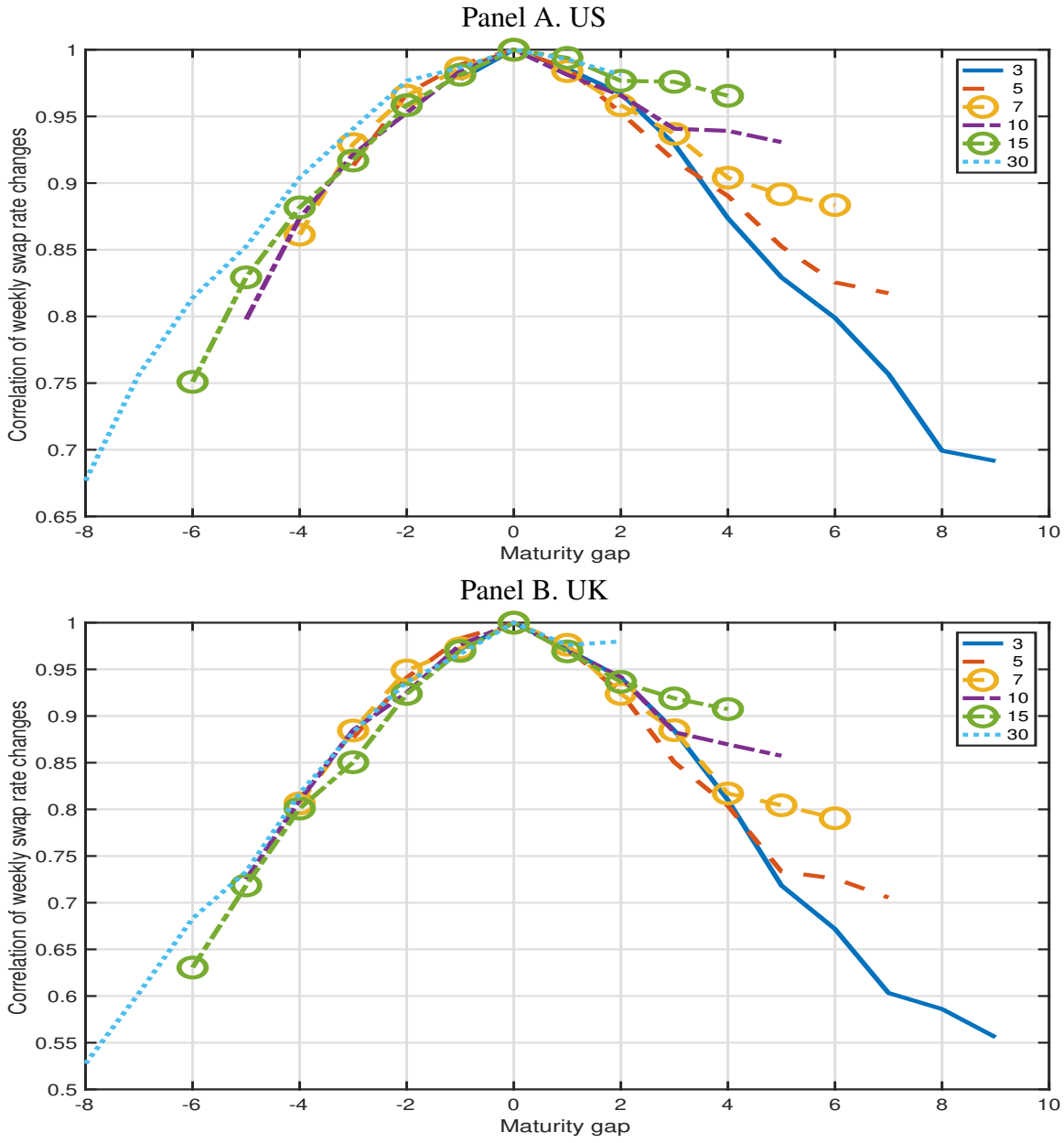


Figure 8

Cross-correlations of weekly changes of swap rates at different maturity gaps

The six lines in each panel denote correlation estimates of weekly swap rate changes between each of the six reference maturities at 3, 5, 7, 10, 15, and 30 years and other maturities. The correlation estimates are plotted against the maturity gap between the other maturities and the reference maturity.