Decomposing Long Bond Returns: A Decentralized Theory*

Peter Carr[1] and Liuren Wu[2]

[1]New York University and [2]Baruch College

Abstract

Classic bond pricing links, i.e., centralizes, bond valuation across all maturities by specifying the dynamics of the short-term interest rate. This paper develops a decentralized theory that prices each bond based purely on the near-term behavior of the bond's own yield. The theory levers the domain expertise of an investor on a particular bond and allows the investor to make pricing and investment analysis on the bond without the shackles of an ambitious centralizing mandate. The theory decomposes the short-term return on a bond with respect to the variation of its own yield. Imposing no dynamic arbitrage on the return decomposition leads to a simple pricing equation relating the bond yield to the market pricing and conditional mean and variance forecasts of the yield's near-term change. The paper illustrates the theory's applications in decentralized investment of a single bond and in the construction and investment of decentralized butterfly bond portfolios.

JEL classification: C13, C51, G12, G13

Keywords: Bond return decomposition; Yield decomposition; Duration; Convexity; Carry effect; Expectation; Risk premium; Local commonality; Butterfly trades

*The authors thank Joel Peress (the Editor), a second editor, two anonymous referees, Zhenyu Cui, Ionut Florescu, Nicola Fusari, Francis Longstaff, Scott Joslin, Joseph Santos, and seminar participants at Baruch College, City University of New York, the Financial Engineering Seminar at Stevens Institute of Technology, the Dykhouse Proram Workshop Series in Money, Banking, and Finance at South Dakota State University, the 2016 JHU-AQR 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. Correspondence: Liuren Wu, Department of Economics and Finance, Baruch College, One Bernard Baruch Way, Box B10-225, New York, NY 10010, USA; email: liuren.wu@baruch.cuny.edu

1. Introduction

Classic bond pricing centralizes bond valuation across all maturities by specifying the full dynamics of the short-term interest rate (e.g., Duffie and Kan (1996)). This centralization has played important roles in practical applications such as information consolidation, dimension reduction, and identifying relative value opportunities along the yield curve. In this paper, we develop a decentralized theory that deliberately circumvents the classic centralizing link between the pricing of a bond and the dynamics and pricing of the short rate. Instead, the theory values each bond on its own, based purely on the current market pricing assumption and the conditional mean and variance forecasts on the near-term changes of the yield of this particular bond.

Compared to classic centralizing efforts, the decentralized theory focuses on a much smaller task, but strives to lever the domain expertise of an investor on a particular bond to generate sharper pricing insights on that bond. In principle, it is much easier to forecast the near-term behavior of the yield of a particular bond that the investor has been tracking, than to predict how an unobservable short rate moves over the next few decades. As with the new trend of decentralization in technology, finance, and other areas, the decentralized bond pricing theory provides a framework for investors to make much more focused pricing analysis and investment decisions on a particular bond without the burden and the shackles of an ambitious centralizing mandate.

Under the classic centralizing framework, the values of bonds across all maturities are made to vary with a common set of interest-rate factors. By comparison, the decentralized theory summarizes the risk of each bond through the near-term variation of its own yield to maturity. It captures the bond's risk exposures through classic duration and convexity definitions with respect to its own yield. The theory decomposes the investment return on a bond into three terms, all linked to the yield of this particular bond. The first term is the carry effect. The investor earns the yield to maturity over the next instant if the yield does not change over time. The second term is the duration effect capturing the return contribution due to the linear change of the yield. The last term is the convexity effect capturing the return contribution due to the nonlinear change in the yield.

The bond return decomposition is completely decentralized to its own yield to maturity, with no direct linkages to the short rate dynamics or the dynamics of any common interest-rate factors. Taking expectation on this decentralized return decomposition under the risk-neutral measure and imposing the condition of no dynamic arbitrage leads to a simple pricing relation on the bond's yield. The pricing relation decomposes the bond yield spread over the short rate into three components, driven respectively by the risk premium, the conditional mean, and the conditional variance forecasts of the near-term changes of the yield. The three components rely on three conditional forecasts of the yield's near-term behavior. The pricing of the bond at any given point in time is determined by these three conditional forecasts at that time.

Since the yield on each bond can be analyzed on its own, there is no pricing error contagion effect from one bond to another, which can happen when estimating a centralized model. Since the current pricing of the bond yield depends on the current mean, variance, and market pricing forecasts on its own near-term change, one does not need to forecast beyond the near term, or to make any assumptions on the long-run dynamics of this yield or any other interest-rate series. Since the sources and dynamics of the conditional forecasts have no bearing on the pricing, one can obtain these forecasts either from separate statistical model estimation or outside expert survey, greatly facilitating cross-platform collaboration across teams with different domain expertise. Investors can also update these forecasts over time in whichever way they see fit. Revising the forecasts, or the underlying forecasting model, does not bring any consistency issue experienced in the recalibration of centralized models.

We illustrate applications of the new theory from the perspective of making short-run investment decisions, first on a single par bond and then on a butterfly bond portfolio. To predict the excess return on a par bond, we apply the decentralized theory to the yield to maturity of this par bond. As the theory decomposes the bond yield into a risk premium component and the contributions from the conditional mean and variance forecasts of the yield change, we can apply the theory to extract the bond's risk premium if we can obtain conditional mean and variance forecasts on its yield change. For illustration, we make simplifying assumptions to obtain the conditional forecasts and examine the predictive power of the extracted risk premium on future excess returns of the par bond. Historical analysis of US long-term swap rates, which we treat as the coupon rates for par bonds, shows that the risk premium extracted from each long-term par bond according to the decentralized theory can be used to generate robust out-of-sample predictions on the bond's future excess return.

Investors often combine bonds of different maturities to hedge away systematic interest-rate movements. Constructing butterfly portfolios with three bonds has been a staple trade in the fixed income market for this purpose. By hedging away major movements in interest rate levels and possibly slope changes, a well-constructed butterfly portfolio can become very stable, allowing investors to achieve targeted exposures to the curvature of the yield curve and potential temporary mispricings on the particular bonds. Historically, dynamic term structure models have been used to construct butterfly portfolios for statistical arbitrage trading based on model pricing errors (e.g., Duarte, Longstaff, and Yu (2007) and Bali, Heidari, and Wu (2009)). Based on the new decentralized theory, we propose to construct decentralized butterfly portfolios with bonds of nearby maturities, and we propose new investment strategies for these decentralized butterflies based on classic risk-return tradeoff analysis of the bond portfolio's excess returns. We show that a combination of shrewd maturity choice for the butterfly portfolio construction and careful mean-variance analysis on the portfolio returns leads to robust investment performance.

Representing the relative value of a bond in terms of its yield to maturity and measuring its risk exposures via the bond's duration and convexity have long been the industry standard (e.g., Christensen and Sorensen (1994) and Chance and Jordan (1996)) and also the starting point of most fixed income textbooks. This paper builds a new pricing theory on a classic foundation by deriving the pricing implications from classic bond return attribution analysis with respect to its own yield to maturity. The new theory integrates no-arbitrage pricing naturally into classic risk management, return attribution, and risk-return investment analysis.

The remainder of the paper is organized as follows. Section 2 develops the new pricing theory. Section 3 illustrates its application in a decentralized investment analysis of a par bond. Section 4 illustrates its application in the construction and investment of decentralized butterfly bond portfolios. Section 5 concludes. An online appendix discusses our data choice for the empirical analysis, reports the summary behaviors of the US swap rates, and performs comparative analysis with stripped US Treasury par yields.

2. 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 the usual conditions of right continuity and completeness with respect to the null sets of $\mathbb{P}$. We further assume the existence of a money market account associated with an instantaneous interest rate (or "short rate" for short) $r_t \geq 0$. The assumption of no dynamic arbitrage implies the existence of an equivalent martingale measure $\mathbb{Q}$, often labeled as the risk-neutral measure, associated to this money market account as the numeraire.

Let B_t denote the time-t value of a riskfree bond with a stream of N cash payments $\{C_j\}_{j=1}^N$ 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. Classic dynamic term structure models start by modeling the dynamics of the short rate r_t and value the bond via the following expectation operations over the horizon of the cash flows,

$$B_t = \sum_{j=1}^{N} C_j \mathbb{E}_t^{\mathbb{P}} \left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) e^{-\int_t^{t+\tau_j} r_u du} \right] = \sum_{j=1}^{N} C_j \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^{t+\tau_j} r_u du} \right]. \tag{1}$$

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, and $\frac{d\mathbb{Q}}{d\mathbb{P}}$ defines the measure change from $\mathbb{P}$ to $\mathbb{Q}$ via the specification of the market pricing of risks. Through the expectation operation, bonds with cash flows at all times are chained together via either the centralized modeling of the short rate dynamics and its market pricing of risk, or directly the risk-neutral short rate dynamics. In particular, valuing a long-dated bond necessitates projections of the short rate dynamics and its pricing over the long horizon of the bond's cash flows.

2.1 Decentralized Carry-Duration-Convexity Bond Return Decomposition

An investor can invest in a very long-dated bond for a very short period of time. In this case, the investor worries more about the short-term value fluctuation of the bond in question than about long-term projections of the short rate. Even for investors with a long investment horizon, managing the daily return fluctuation remains vitally important. Given these practical considerations, our new pricing theory does not rely on the long-term projection of a short rate, but builds on the short-term return attribution of the particular bond investment.

To decentralize the return attribution, we represent the bond value not as a function of the short rate or some common interest-rate factors, but in terms the bond's own yield to maturity. Given the price of a particular bond B_t, its yield to maturity y_t is defined via the following equality,

$$B_t \equiv \sum_{j=1}^{N} \exp\left(-y_t \tau_j\right) C_j. \tag{2}$$

The yield to maturity is decentralized in the sense that it is unique to the bond in question. Given the set of cash flows for the bond, the yield to maturity of the bond has a unique and monotone mapping with the price of the bond. Indeed, it has become standard industry practice to quote the price of a bond in terms of its yield to maturity.

With the decentralized yield to maturity as defined in (2) to summarize a bond's risk exposure, we attribute the short-term return of the bond investment with respect to the variation of its own yield to maturity as,

$$\frac{dB_t}{B_t} = \frac{\partial B_t}{B_t \partial t} dt + \frac{\partial B_t}{B_t \partial y_t} dy_t + \frac{1}{2} \frac{\partial^2 B_t}{B_t \partial y_t^2} (dy_t)^2 + o(dt), \tag{3}$$

where $o(dt)$ denotes higher-order terms of dt when the 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 is continuous, and attribute the return solely to the time effect and the first and second-order effects from the yield movement. Since the attribution focuses on the bond return over the next instant, the continuity assumption applies only for the next instant. The results hold even if the yield can jump at any other times.

From the yield to maturity definition in (2), we can derive the three sensitivity terms in (3) as

$$\frac{\partial B_t}{B_t \partial t} = y_t, \quad -\frac{\partial B_t}{B_t \partial y_t} = \sum_{j=1}^{N} w_j \tau_j, \quad \frac{\partial^2 B_t}{B_t \partial y_t^2} = \sum_{j=1}^{N} w_j \tau_j^2, \tag{4}$$

where the weight w_j on each cash flow j reflects its present value as a fraction of the total value,

$$w_j = \frac{\exp\left(-y_t \tau_j\right) C_j}{\sum_{i=1}^{N} \exp\left(-y_t \tau_i\right) C_i}. \tag{5}$$

The first term of the bond return decomposition in (3) measures the carry effect of the investment: Carrying the bond makes y_t per year if the yield stays constant. The second term measures the directional exposure of the bond investment to the yield to maturity variation. Bond price declines as the yield to maturity increases and the sensitivity is determined by the weighted average maturity, or *duration*, of the bond. The third term captures the nonlinearity, or *convexity*, of the price-yield relation. The positively convex relation increases the bond return when the yield varies more. We use $\mathcal{D}$ and $\mathcal{C}$ to denote the duration and convexity of the bond, respectively,

$$\mathcal{D} \equiv \sum_{j=1}^{N} w_j \tau_j, \quad \mathcal{C} \equiv \sum_{j=1}^{N} w_j \tau_j^2. \tag{6}$$

For a zero-coupon bond, the thus-defined duration is equal to the time to maturity of the bond τ, and the convexity is equal to the squared maturity τ^2. For coupon bonds, the duration and convexity are the value-weighted averages of the maturities and squared maturities, respectively.

Theorem 1 (Decentralized Carry-Duration-Convexity Bond Return Decomposition)**.** *Under the assumption of diffusive yield moves over the next instant and with the decentralized definition of duration and convexity in (6) with respect to the bond's own yield to maturity, the short-term return of a bond investment can be decomposed into carry, duration, and convexity exposures, respectively,*

$$\frac{dB_t}{B_t} = y_t dt - \mathcal{D} dy_t + \frac{1}{2}\mathcal{C}(dy_t)^2. \tag{7}$$

Taking expectation on (7) under the statistical measure $\mathbb{P}$ and dividing both sides by the time interval dt attributes the annualized expected bond investment return to three sources,

$$\frac{1}{dt}\mathbb{E}_t^{\mathbb{P}}\left[\frac{dB_t}{B_t}\right] = y_t - \mu_t \mathcal{D} + \frac{1}{2}\sigma_t^2 \mathcal{C}, \tag{8}$$

where $\mu_t = \mathbb{E}_t^{\mathbb{P}}[dy_t]/dt$ denotes the time-t expected rate of change on the yield, and $\sigma_t^2 = \mathbb{E}_t\left[(dy)^2\right]/dt$ denotes the time-t conditional variance rate of the yield change. 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, an expected increase in yield reduces the expected bond return. Third, higher volatility on the yield movements leads to higher expected bond return due to the positive convexity effect.

2.2 Decentralized No-arbitrage Pricing and Yield Decomposition

To generate pricing implications, we take expectation under the risk-neutral measure $\mathbb{Q}$ on the bond return attribution in (7) and divide both sides by dt,

$$\frac{1}{dt}\mathbb{E}_t^{\mathbb{Q}}\left[\frac{dB_t}{B_t}\right] = y_t - \mu_t^{\mathbb{Q}} \mathcal{D} + \frac{1}{2}\sigma_t^2 \mathcal{C}, \tag{9}$$

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. If we use λ_t to denote the market pricing per unit of the bond risk,[1] under the locally diffusive movement assumption, we can link the risk-neutral expected rate of yield change $\mu_t^{\mathbb{Q}}$ to its statistical counterpart μ_t by

$$\mu_t^{\mathbb{Q}} = \lambda_t \sigma_t + \mu_t. \tag{10}$$

The market price is positive $\lambda_t > 0$ if bond returns are expected to contain a positive risk premium.

The absence of dynamic arbitrage between this particular bond under consideration and the money market account dictates that the risk-neutral expected rate of return on the bond is equal to

[1] Given the inverse price-yield relation, the market price of yield risk takes the opposite sign of the market price of bond risk.

the short ate r_t. Applying this condition of no dynamic arbitrage to the risk-neutral expectation in (9) leads to a simple pricing relation for the bond under consideration.

Theorem 2 (Decentralized Yield Decomposition)**.** *If the yield of a bond is moving continuously over the next instant, no dynamic arbitrage between this bond and the money market account dictates that the fair spread of this bond yield over the short rate is linked to the market pricing of this bond's risk (λ_t), the expected rate of the yield change (μ_t), and its variance rate (σ_t^2) through the bond's duration $\mathcal{D}$ and convexity $\mathcal{C}$ exposures by*

$$y_t - r_t = \lambda_t \sigma_t \mathcal{D} + \mu_t \mathcal{D} - \frac{1}{2}\sigma_t^2 \mathcal{C}. \tag{11}$$

The pricing relation in (11) is highly decentralized. The fair value of the bond's yield spread $(y_t - r_t)$ is determined by the market pricing of this particular bond (λ_t) and the conditional mean (μ_t) and variance (σ_t^2) forecasts on the next movement of the bond's own yield. The valuation has no direct dependence on the short rate dynamics or the dynamics of any other yields. Indeed, there should be an indexing i on $(y_t, \mathcal{D}, \mathcal{C}, \lambda_t, \mu_t, \sigma_t^2)$ to emphasize that they are all specific to the particular bond i under consideration. The valuations of two different bonds, e.g., i and j, are based on two separate sets of market pricing assumptions and conditional mean and variance forecasts. We omit this indexing to reduce notation cluttering when no confusion shall occur.

Even for the particular bond under consideration, the time-t pricing of the bond does not rely on the full dynamics of its own yield, but only depends on the three point-in-time conditional forecasts at that time $(\mu_t, \sigma_t, \lambda_t)$. The three conditional forecasts can change over time, but how they change, or where the forecasts come from, does not enter the pricing of the current yield spread. One can bring in the forecasts from any outside sources and directly examine their pricing implications under this decentralized theory.

2.3 Comparison with Centralized Term Structure Modeling

The pricing relation in (11) is obtained by attributing the bond return to variations in its own yield. Retaining the partial derivatives in the return attribution, we can write the pricing equation as a partial differential equation,

$$r_t = \frac{\partial B_t}{B_t \partial t} + \frac{\partial B_t}{B_t \partial y}\mu_t^{\mathbb{Q}} + \frac{1}{2}\frac{\partial^2 B_t}{B_t \partial y^2}\sigma_t^2, \tag{12}$$

which is the result of imposing the no dynamic arbitrage condition that the risk-neutral expected rate of return on the bond investment is equal to the short rate r_t.

Dynamic term structure models value bonds by imposing the same no dynamic arbitrage condition; however, the centralized modeling approach does not attribute bond returns with respect to its own yield to maturity, but rather with respect to a set of common factors X_t that determine the dynamics of the short rate r_t. Assuming diffusive dynamics for the common factor X_t with μ_t^X denoting its risk-neutral drift vector and Σ_t^X its covariance matrix, we can express the same no dynamic arbitrage condition via an analogous partial differential equation,

$$r(X_t) = \frac{\partial B_t}{B_t \partial t} + \left[\frac{\partial B_t}{B_t \partial X_t}\right]^\top \mu_t^X + \frac{1}{2} tr\left[\frac{\partial^2 B_t}{B_t \partial X X^\top} \cdot \Sigma_t^X\right]. \tag{13}$$

The two partial differential equations in (12) and (13) are similar in form and represent the same no dynamic arbitrage condition, but they have crucially different implications. First, the relation between the bond price and its own yield is explicit and definitional. As a result, the partial derivatives in (12) are known ex ante. Plugging in these partial derivatives yields the simple algebraic pricing relation in (11). By contrast, the bond price as a function of the common factor vector, $B(X_t)$, is unknown ex ante. The partial differential equation in (13) is used to solve for this pricing relation, starting with the boundary conditions at the cash payments dates of the bond.

Second, to solve the partial differential equation in (13) backward starting at a bond's payment date T, one must necessarily fully specify the common factor dynamics (μ_s^X, Σ_s^X) across all times from today t to the payment date T. By comparison, the decentralized pricing relation only needs the current mean and variance forecasts of its yield change for its current pricing, without the need to know its future dynamics.

Finally, since the short rate is represented as a function of the common factors in (13), once the common factor dynamics are specified, the values of bonds of all maturities are set, true to the centralizing nature of the modeling approach. By contrast, knowing the conditional mean and variance forecasts of the yield changes for one bond at time t determines the time-t valuation of the yield for this particular bond, but it does not pin down the pricing of the bond at any other times, nor the pricing of any other bonds, true to the decentralized nature of the new pricing approach. Therefore, the partial differential equation in (12) only represents a decentralized and point-in-time no arbitrage condition at a particular time t between the particular bond under consideration and the money market account, whereas the partial different equation in (13) represents a centralized no arbitrage condition between all bonds and the money market account at any time t.

In principle, one can centralize the new pricing relation by centralizing the mean and variance forecasts on yield changes across different maturities via a common factor structure, but a dynamic term structure model provides a more natural way of achieving this centralization. On the other hand, bond valuation derived from a dynamic term structure model naturally satisfies the new pricing relation. The model-implied yield level must be consistent with the model-implied conditional mean and variance of the corresponding yield change. Nevertheless, under the new decentralized theory, investors can directly make assumptions and/or statistical estimates on the conditional mean and variance of a yield change, and derive implications on the fair valuation of the particular bond, without the need to make assumptions on any long-run dynamics of the short rate. Therefore, the two modeling approaches do not directly compete, but rather complement each other by starting from different perspectives.

Through centralization, dynamic term structure models can readily achieve its main objective of imposing cross-sectional consistency on the valuation of different bonds. The consistency is achieved by using the same set of common factor dynamics to price bonds of all maturities. The achievement, however, does not come for free. First, it is not easy to specify one common dynamics that matches the behavior of all bonds. As one considers the pricing of more bonds across a wider spectrum of maturities, one is often forced to expand the dimensionality of the common factor structure to accommodate the wider range of behaviors. A model that looks sufficient for a subset of data can generate gross mispricing when the data set is expanded. Second, centralization is good for information consolidation, dimension reduction, and noisy reduction, but it can also generate error contagion across all bonds with the introduction of a data error on a single bond.

Through decentralization, the new pricing theory cannot tell us much about cross-sectional consistency. Each bond is priced to be consistent with its own mean and variance forecasts, but there is no direct way of checking whether the forecasts on one bond are consistent with the forecasts on another bond. On the other hand, the decentralization allows investors to better lever one's domain expertise on a particular set of bonds to generate better pricing on these bonds based on more narrowly-focused but potentially more insightful forecasts. Jack of all trades, master of none. As one is unlikely to become the master of everything, the decentralized theory allows one to aim small and miss small.

The different perspectives also lead to different starting points and accordingly different assumptions. To do centralized pricing according to (13), one must specify the full dynamics of the short rate. Since our knowledge about the far distant future is limited, researchers often rely on stationarity assumptions to generate limiting behaviors in the long run. Seemingly innocuous stationarity assumptions, however, can generate grossly violated extrapolation results. For example, based on stationarity assumptions, Dybvig, Ingersoll, and Ross (1996) make the bold prediction that long-run forward rates can never fall in the limit. While the prediction is consistent with the assumption and the limit can always be set longer than any observation, the reality is that the observed ultra long-dated interest rates up to 50 years vary as much as rates at shorter maturities. Forward rates at these observable ultra long maturities tend to fall lower than forward rates at intermediate maturities.

In a comparative analysis across several markets, Giglio and Kelly (2018) show that mean-reverting dynamics calibrated to short-dated securities often generate variations that are too small for long-dated securities. As bonds of increasingly long maturities are becoming available, researchers are becoming increasingly aware of the practical constraints of the long-run stationarity assumptions and have been working hard to find remedies (Bauer and Rudebusch (2020)). The decentralized theory does not require one to reconcile the current behavior of long-dated yields with any long-run assumptions of the short rate dynamics. Instead, investors can take the recent observed behaviors of the bond's yield as the starting point and take the recent volatility estimates on a bond's yield changes as direct inputs to make pricing analysis and investment decisions under the decentralized theory.

3. Decentralized Investment in A Single Bond

The decentralized theory in (11) decomposes the yield spread on a particular bond at any given point in time ($y_t - r_t$) into contributions from the market pricing of this bond's risk (λ_t) and the conditional mean (μ_t) and variance (σ_t^2) forecasts on the next movement of the bond's yield. Rearrange the pricing relation, we can represent the market pricing of the bond's risk in terms of the observed yield spread and the conditional mean and variance forecasts,

$$\lambda_t = \frac{y_t - r_t - \mu_t \mathcal{D} + \frac{1}{2}\sigma_t^2 \mathcal{C}}{\sigma_t \mathcal{D}}. \tag{14}$$

The pricing of a bond as represented by its yield spread $y_t - r_t$ can vary either because of variations in the conditional mean and variance forecasts of the yield change (μ_t, σ_t^2), or because of variations in the market pricing of risk (λ_t). With the risk exposures ($\mathcal{D}, \mathcal{C}$) fixed, when the expectations do not change, the variation of the bond price (or equivalently its yield) purely reflects variation in the market pricing of risk.

Therefore, once an investor can formulate the conditional mean and variance forecasts on the yield change of a bond to fix the expectation, the investor will be able to determine the market pricing of the bond's risk based on the observed yield spread on the bond and the bond's duration and convexity estimates. The market pricing of the bond's risk, combined with the risk forecast (σ_t), determines the expected excess return, or risk premium, that the investor expects to earn from the bond investment. In this section, we illustrate how an investor can apply the theory to extract the market pricing of a bond's risk based on the observed yield of the bond and a set of conditional mean and variance forecasts on its yield change.

3.1 Extracting Market Price of Bond Risk from its Own Yield Behavior

To extract the market price of a bond's risk, the theory requires as inputs the conditional mean and variance forecasts on the change of the bond's yield. As an illustration, we make stylized assumptions to obtain the conditional forecasts and extract the market price of bond price based on these forecasts. We discuss potential extensions to generate enhanced conditional forecasts in a later subsection.

To construct the conditional mean forecast μ_t, we note that long-term constant-maturity floating interest-rate series are highly persistent and changes in these floating series are extremely difficult to predict. We take this difficult-to-predict nature of long-term floating rates as our starting point. Specifically, by assuming that the expected rate of change of the constant-maturity yield at $\tau = T - t$ is zero, we can predict the expected rate of change of the yield on the bond with fixed expiry T based on the yield's sliding effect along the yield curve,

$$\mu_t = -s_t, \tag{15}$$

with s_t denoting the local slope of the yield curve around the bond's time to maturity τ. Intuitively, if the yield curve is upward sloping and is not expected to change over time, as time t passes by, the time to maturity $\tau = T - t$ of the bond with a fixed expiry T will shrink, and the bond's yield will become lower as it slides down the upward sloping yield curve with declining time to maturity. The rate of decline on the bond yield is determined by the local slope of the yield curve around the bond's maturity. We propose to estimate the local slope of the yield curve around the bond's maturity with a local linear regression and construct the conditional mean forecast based on the local slope estimate $\widehat{s_t}$.

To construct the conditional variance forecast on the bond's yield, we assume that the yield's risk behavior will be similar to the recent historical behavior of the yields at similar maturities.

Accordingly, we propose to construct a historical variance estimator on the yield changes $\widehat{\sigma}_t^2$ and use the estimator as the conditional variance forecast .

Combining the observed yield spread $y_t - r_t$, its duration $\mathcal{D}$ and convexity $\mathcal{C}$ exposures, the local yield curve slope estimate $\widehat{s}_t$, and the historical variance estimator $\widehat{\sigma}_t^2$, we can infer the market price of the bond's risk.

Proposition 1. *Assuming no-predictability on long-term constant-maturity yields and using historical variance estimator as a simple variance forecast, one can estimate the time-t market price of the risk on a long-term bond, λ_t, from the time-t observed yield spread $y_t - r_t$, its duration $\mathcal{D}$ and convexity $\mathcal{C}$ exposures, the local yield curve slope estimate $\widehat{s}_t$, and the historical variance estimator $\widehat{\sigma}_t^2$ of the yield change on this bond as,*

$$\widehat{\lambda}_t = \frac{y_t - r_t + \widehat{s}_t \mathcal{D} + \frac{1}{2}\widehat{\sigma}_t^2 \mathcal{C}}{\widehat{\sigma}_t \mathcal{D}}. \tag{16}$$

We perform empirical analysis on the Proposition using US swap rates. We treat the swap rates as the coupon rates of par bonds and use the financing leg of the swap contract (LIBOR) as the short rate r_t. We obtain the LIBOR and swap rate data from Bloomberg weekly every Wednesday from January 4th, 1995 to December 26, 2018, for 1,252 weeks. The data include nine swap maturities at 2, 3, 4, 5, 7, 10, 15, 20, and 30 years. An online appendix discusses the data choice, the summary behaviors of the swap rates, and a comparative analysis on stripped US Treasury par yields.

We estimate the local swap rate curve slope $\widehat{s}_t$ at each maturity using a local linear regression of the swap rates against the time to maturity. We construct historical volatility estimators on weekly changes of each swap rate series with a one-year rolling window, and we further smooth out noises in the individual variance estimators by performing a local linear smoothing regression of the estimators across maturity. We estimate the duration and convexity of each par bond as the value-weighted maturity and squared maturity of the bond's coupons and principals. We strip the swap rate curve by assuming piece-wise constant forward rates and value the coupons and principals based on the stripped discount rate curve.

Figure 1 plots the time series of the market price of bond risk ($\widehat{\lambda}_t$) extracted from selected long-term swap rate series. The market prices of bond risk extracted from different long-dated swap rate series are similar in magnitude and move closely together. The cross-correlation estimates between the different $\widehat{\lambda}_t$ series with maturities 10 years and longer average at 99.7%. Similar to findings in Cochrane and Piazzesi (2005), our evidence supports the existence of a common market price of risk factor for bonds across different maturities.

[Figure 1 about here.]

On average, the market price of bond risk estimates are positive. Nevertheless, the estimates vary strongly over time. In particular, the estimates went below zero right before the start of the two recessions in 2000 and 2007, respectively. At the end of our sample in December 2018, the market price estimates dip below zero again, foretelling the possible incoming of another recession.

3.2 Predicting Bond Excess Returns with Bond Risk Premium Estimators

Given the extracted market price of risk $\widehat{\lambda}_{i,t}$ on each long-term par bond i and the corresponding volatility estimator $\widehat{\sigma}_{i,t}$, we can construct a bond risk premium estimator as

$$\widehat{CW}_{i,t} = \widehat{\lambda}_{i,t}\,\widehat{\sigma}_{i,t}\,. \tag{17}$$

In theory, the annualized expected instantaneous excess return on the par bond is proportional to the bond risk premium and the bond's duration,

$$\frac{1}{dt}\mathbb{E}_t^{\mathbb{P}}\left[\frac{dB_{i,t}}{B_{i,t}}\right] - r_t = \lambda_{i,t}\sigma_{i,t}\mathcal{D}_{i,t}\,. \tag{18}$$

We test this theoretical implication by examining whether our bond risk premium estimator ($\widehat{CW}_{i,t}$) has any actual predictive power on the future bond excess returns. We compute six-month ahead excess return on each par bond based on the stripped discount curve six months later. The choice of the half-year horizon matches the semi-annual coupon payment and the financing rate horizon.

Since the volatility of yield changes is similar across maturities, the variation of the bond excess returns increases with the bond duration $\mathcal{D}_{i,t}$. We annualized the excess return on each par bond and divide it by the bond duration to make the magnitude of the scaled excess returns more comparable across maturities. We examine the predictability of the bond risk premium via the following proportional forecasting regression,

$$ER_{i,t+h} = c_i \widehat{CW}_{i,t} + e_{i,t+h}, \tag{19}$$

where $ER_{i,t+h}$ denotes the annualized and duration-scaled excess return on the ith par bond over the half-year future horizon. The regression allows a proportional coefficient to accommodate scaling differences between the risk premium estimator $\widehat{CW}_{i,t}$ and the bond excess return $ER_{i,t+h}$, but it does not include an intercept term, essentially imposing the constraint that when the risk premium estimator is zero, so will be the expected bond excess return.

For comparison, we consider as a benchmark the bond risk premium estimator constructed by Cochrane and Piazzesi (2005) with a portfolio of forward rates with maturities from 1 to 5 years. To construct this alternative bond risk premium estimator, we strip the forward rate curve from the observed swap rates, and estimate the weight on the forward rate portfolio by performing a forecasting regression of the average future excess returns on the five forward rates,

$$AER_{t+h} = \beta_0 + \sum_{j=1}^{5} \beta_j f_t^j + e_{t+h}, \tag{20}$$

where $AER_{t+h} = \frac{1}{9}\sum_{i=1}^{9} ER_{i,t+h}$ denotes the average excess returns on the nine par bond series that we have from maturities 2 to 30 years. The common bond risk premium factor is constructed from the estimated relation as,

$$\widehat{CP}_t = \widehat{\beta}_0 + \sum_{j=1}^{5} \widehat{\beta}_j f_t^j. \tag{21}$$

With the common risk premium estimator $\widehat{CP}_t$, we forecast the excess return on each par bond via an analogous proportional forecasting regression,

$$ER_{i,t+h} = c_i \widehat{CP}_t + e_{i,t+h}. \tag{22}$$

We examine the predictability of the two bond risk premium estimators through an out-of-sample exercise. Starting from January 5, 2005, on each date t, we estimate the forecasting relations from (19) to (22) with a 10-year rolling window. We compute the out-of-sample forecasting error on each date t as the difference between the future realized and the forecasted excess return,

$$e_{i,k,t+h} = ER_{i,t+h} - \widehat{ER}_{i,k,t}, \tag{23}$$

where $\widehat{ER}_{i,k,t}$ denotes the time-t out-of-sample forecast from predictor k on excess return $ER_{i,t+h}$. As in Welch and Goyal (2008), the forecasting error of each predictor is compared with the forecasting error using the 10-year rolling window historical average of the bond excess return up to that point t,

$$e_{i,0,t+h} = ER_{i,t+h} - \overline{ER}_{i,t}, \quad \overline{ER}_{i,t} = \frac{1}{L}\sum_{s=t-L-h}^{t-h} ER_{i,s+h}, \tag{24}$$

with L denoting the rolling window length. We measure the out-of-sample performance of each predictor via an out-of-sample R^2 measure as in Rapach and Zhou (2013),

$$R^2_{i,k} = 1 - \frac{\sum_{t=1}^{M} e^2_{i,k,t+h}}{\sum_{t=1}^{M} e^2_{i,0,t+h}}, \tag{25}$$

with M denoting the number of out-of-sample observations. A positive R^2 estimate indicates that the predictor outperforms the historical average benchmark.

To test the statistical significance of the forecasting performance difference over the historical average benchmark, we compute the Diebold and Mariano (1995) t-statistics (DM) on the squared forecasting error difference, $\delta_{k,t+h} = e^2_{i,0,t+h} - e^2_{i,k,t+h}$,

$$\mathrm{DM}_k = \frac{\mu_\delta}{\sigma_\delta}\left(\frac{M+1-2(h+1)+h(h+1)/M}{M}\right)^{0.5}, \tag{26}$$

where μ_δ denotes the sample mean of the difference, σ_δ denotes the Newey and West (1987) standard error estimate, computed with a lag equal to the forecasting horizon. The statistics adjust for the small sample size bias according to Harvey, Leybourne, and Newbold (1997). Under the null hypothesis that each rolling-window estimated model and the historical average benchmark have equal finite-sample forecast accuracy, Clark and McCracken (2012) find that the thus-computed DM test statistic can be compared to standard normal critical values.

Table I reports the out-of-sample R^2 estimates and the DM test statistics against the historical average benchmark for each par bond. Panel A reports the performance of the benchmark Cochrane-Piazzesi (CP) bond risk premium estimator constructed with a portfolio of forward rates. Panel B reports the performance of our (CW) bond risk premium estimators based on a simple implementation of our decentralized theory. The out-of-sample statistics are computed over 704 weekly observations from January 5th, 2005 to Jun 27th, 2018 for excess returns over the next six months.

[Table I about here.]

Our no-predictability assumption on floating interest rates in generating the conditional mean forecast is likely to hold better for long-term swap rates than for short-term swap rates, the movements of which can in principle be predicted by the slope of the interest-rate term structure. On the other hand, the Cochrane and Piazzesi (2005) forward-rate portfolio was originally constructed for predicting excess returns on short-term bonds with maturities from 2 to 5 years. Table I reports the forecasting performance of the two risk premium estimators for both short-term and long-term par bonds.

Panel A shows that the CP bond risk premium estimator generates negative out-of-sample forecasting R^2 estimates across all maturities. By contrast, Panel B shows that our CW bond risk premium estimators generate positive out-of-sample forecasting R^2 estimates across all maturities. Furthermore, the DM statistics show that the out-of-sample forecasting performance of the CW bond risk premium estimator is significantly better than the historical average benchmark for long-term bonds with maturities from 5 to 30 years. As a robustness check, the online appendix repeats the analysis on stripped US Treasury par yields over a longer sample period. The findings are qualitatively similar.

3.3 Discussions

The decentralized theory separates the risk premium component from the observed yield spread on a bond based on the conditional mean and variance forecasts on the next movement of the bond's yield. This section illustrates the application of the theory in extracting the bond risk premium by making stylized assumptions in constructing the conditional mean and variance forecasts. For future research, with robust estimation methodologies that can mitigate in-sample overfitting and out-of-sample deterioration, researchers can incorporate more information to generate better conditional mean and variance forecasts. Better conditional forecasts should lead to a better separation of the bond risk premium from the observed yield spread. The information for the conditional forecasts can come from historical behaviors of the yield changes, forward-looking information in interest-rate options, the shape of the yield curve, macroeconomic variables, and surveys of expert opinions (e.g., Kim and Orphanides (2012)). The conditional forecasts can be constructed via either statistical specifications or some structural model estimation. The decentralized theory can readily incorporate these conditional forecasts to generate a risk premium estimator, without the need to know the source or the underlying mechanism of the forecasts.

For the general purpose of bond return prediction, different modeling framework offers different insights on the potential sources of predictive information. The classic expectation hypothesis literature (e.g., Fama and Bliss (1987) and Campbell and Shiller (1991)) focuses on the information content of the yield curve slope in predicting future interest rate changes and in making inferences about bond risk premium behaviors. Cochrane and Piazzesi (2005) propose to expand beyond the yield curve slope and use a portfolio of five forward rates to capture the potentially richer information content in the shape of the yield curve in predicting bond returns. By contrast, our theory focuses less on the shape of the yield curve, but more on the conditional mean and variance forecasts of the yield change on the bond in question. The theory adjusts the observed yield spread to account for contributions from the expectations in the conditional mean and variance of the bond's yield change to separate out the risk premium component.

In addition to the information in the yield curve shape and the conditional mean and variance forecasts on yield changes, there is also a large strand of macro-finance literature that strives to understand the macroeconomic foundations of the bond risk premium. The literature has identified a long list of macroeconomic variables that show predictability on bond returns. Examples include Cooper and Priestley (2009), who link the bond risk premium to the output gap; Ludvigson and Ng (2009), who extract factors from a large set of macro variables to predict bond returns; Greenwood and Vayanos (2014), who link the Treasury bond return to Treasury bond supply; Joslin, Priebsch, and Singleton (2014), who include measures of economic growth and inflation into dynamic term structure modeling; and Cieslak and Povala (2015), who decompose Treasury yields into long-horizon inflation expectations and maturity related cycles.

While these studies can all be informative about the bond risk premium behavior, how to accurately estimate bond risk premium in an out-of-sample setting remains a challenging task (Gargano, Pettenuzzo, and Timmermann (2017)). Our out-of-sample exercise also highlights this challenge. Over our sample period, the full in-sample bond return forecasting R^2 estimates on the CP bond risk premium estimator are from 14% for the 30-year bond to 32% for the 2-year bond. The highly positive R^2 estimates highlight the potentially rich information content in the shape of the forward rate curve in predicting bond excess returns. By comparison, the in-sample forecasting R^2 estimates from our CW bond risk premium estimators are lower, from 5% for the 30-year bond to 12-14% for bonds at 2-7 years. The lower R^2 estimates suggest that the flexible forward rate curve can in principle capture richer information about future bond excess returns than our risk premium estimators based on simple conditional mean and variance forecasts. Nevertheless, the out-of-sample performance reversal of the two risk premium estimators shows that for out-of-sample return prediction applications, it is important not only to identify informative predictors, but also to have a robust estimation methodology that can reduce in-sample overfitting and out-of-sample deterioration.

Dynamic term structure models can be used to capture the yield curve shape with a few dynamic factors. The associated dimension reduction can in principle reduce error-in-variables issues and accordingly reduce the estimation errors and biases when one uses the extracted yield curve factors to predict future bond returns. The issue is that these factors are extracted to maximize the explanatory power on the variation of the yield curve, but they do not necessarily capture the most informative yield curve information about future bond returns (Duffee (2011)).

Within the context of inflation rate forecasting, Hua and Wu (2018) highlight the out-of-sample instability of predictive regressions and propose ways of predicting inflation and interest rate changes without predictive regressions. Cieslak and Povala (2015) also avoid running predictive regressions by extracting the bond risk premium via a contemporaneous regression of a long-term yield on trend inflation and the short rate. Most recently, Bianchi, Büchner, and Tamoni (2021) show that machine learning techniques can potentially take on the challenge by being able to incorporate a long list of predictive variables, allow nonlinear and non-additive relations, and at the same time maintain out-of-sample stability.

4. Investing in Decentralized Butterfly Bond Portfolios

The previous section shows how investors can apply the decentralized theory to examine the risk-return behavior of a single bond. In this section, we apply the theory to construct stable, decentralized butterfly portfolios with bonds of nearby maturities, and propose new investment strategies on these decentralized butterflies based on classic risk-return tradeoff analysis.

Butterfly portfolios constructed with three bonds have been a staple trade in the fixed income market. By hedging away systematic movements in the interest-rate levels and possibly also slope changes, a well-constructed butterfly portfolio can become very stable, allowing investors to achieve targeted exposures to the curvature of the yield curve and potential temporary mispricings of the bonds in the portfolio. Duarte, Longstaff, and Yu (2007) construct butterfly portfolios based on a two-factor Gaussian affine model and examine their risk-return behaviors. Two-factor Gaussian affine models have been the anchor for many quantitative fixed income desks since the 1980s. A three-leg butterfly portfolio can be constructed to neutralize the two common factors of the model. If the two-factor structure is well-specified and the pricing errors are mean-reverting, the constructed butterfly portfolios can become much more mean-reverting than the individual interest-rate series, and one can treat the negative of the pricing errors as the alpha source. Such a portfolio construction and investment process is typical of a model-based statistical arbitrage trading strategy. The stability and profitability of such a strategy depends crucially on how well the model specification captures the true underlying factor structure and how successful the model can separate persistent interest-rate movement from the more mean-reverting residual movements.

Given the centralized nature of a dynamic term structure model, the portfolio construction is in theory agnostic to the particular bond maturity choice: One can in principle hedge away the two interest-rate factors with combinations of any three distinct interest-rate series. Nevertheless, 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 theory, we propose to construct decentralized butterfly portfolios that do not rely on factor structure assumptions but focus much more on the particular choice of the maturity combination. We propose to invest in these decentralized butterflies, not based on statistical arbitrage trading of the pricing errors of a term structure model, but based on classic risk-return analysis of the portfolio's excess returns.

4.1 Local Commonality of Interest-rate Movements

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

To illustrate this local commonality 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 20 years, respectively. We take these maturities as reference maturities and show how the correlations of weekly changes of the swap rate series between the reference series and series at other maturities decline as their maturity gaps increase.

Figure 2 plots the sample correlation estimates against the maturity gap with the reference maturity. Each line plots the correlation of one reference maturity with other maturities. Our data sample includes weekly swap rate series from January 1995 to December 2018 at nine maturities from two to 30 years. The maturity gap measures the gap between the nine 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 2 about here.]

The correlation estimates decline monotonically as the absolute maturity gap increases. The correlations of weekly changes of swap rates at adjacent maturities (with maturity gaps of ± 1) are all higher than 97.78%, with an average of 98.58%. The estimates remain high at the next level (with gaps of ± 2) at no less than 95.27%, with an average of 96.32%. As the gap increases further, the average correlation estimates decline from an average of 92.7% with gaps of ± 3 to an average of 73.86% with gaps of ± 7.

Local commonality is not an exact statement. It is a qualitative observation of an approximate but robust and universal phenomenon. By comparison, when one specifies a global factor structure, its implications are exact, but the exact implications are not necessarily robust in the sense that the exact implications can be wrong when the model assumptions are violated.

To design global factor structures for the term structure of interest rates, researchers often start with principal component analysis. The findings of such analysis, however, depend crucially on the number of interest-rate series and the maturity span of the series included in the analysis. In general, the more series one includes in the analysis and the wider is the maturity span, the more principal components one finds necessary to explain a sufficiently high percentage of the variation of the term structure. As such, any global factor structure assumptions are likely to be only locally true to the scope of the data under analysis. The only thing that is truly global is the local commonality of interest rates with nearby maturities, regardless of how far the data extend and what the true underlying dynamics are.

The history of the US bond market highlights this fluid nature of a global factor structure. The US bond market in the early years is concentrated in the short maturity segment between two to five years. Accordingly, the historical literature on expectation hypothesis, including the bond risk premium analysis by Cochrane and Piazzesi (2005), focuses on this short end of the term structure, where the convexity effect is negligible and the expectation of future interest-rate movements dominates the term structure shape. A one-factor model, such as the classic Vasicek (1977) model or the Cox, Ingersoll, and Ross (1985) model, can do a reasonably good job explaining the term structure variation within this short maturity range. As the maturity range of the bond market expands to up to 10 years, quantitative researchers in the industry start to implement two-factor term structure models to capture variations over this wider maturity span by, for example, allowing the short rate to revert to a stochastic central tendency instead of a constant mean (Balduzzi, Das, and Foresi (1998)). Butterfly trades are a natural result of such two-factor structures. Longer-term Treasury bonds and swap rates with maturities up to 30 years become more actively traded since the early 1990s. The factor analysis of Litterman and Scheinkman (1991) over this maturity range becomes the watershed evidence for establishing the existence of a three-factor structure. This finding has also become the starting point for subsequent specification analysis of dynamic term structure models (Dai and Singleton (2000)).

During the past decade, as the levels of long-term interest rates become increasingly lower across the world, central governments race to issue ultra long-dated bonds, with maturities extending to 60 years. The corresponding swap market has also started trading contracts with ultra long maturities. Including such ultra long maturities in the analysis is bound to put further stress on a three-factor model structure and ask for even more principal components. In fact, a standard three-factor Gaussian affine model has been shown to become increasingly inadequate in capturing the full term structure variation over a wide maturity span. The inadequacy happens at both the short end and the long end. At the short end, Fed policy inertia (Woodford (1999)) reduces the short-rate volatility and induces a time-inhomogeneous trend component that is difficult to be accommodated with any time-homogeneous factor dynamics assumptions (Heidari and Wu (2010)). Short rate hitting the zero lower bound creates another layer of complication and tension for standard affine term structure models, potentially asking for option-like modeling approaches (e.g., Black (1995) and Krippner (2012)). At the very long end, 50-year rates move with as much volatility as interest rates at 10 to 30-year maturities, adding tension to the standard short-rate stationarity assumptions (Bauer and Rudebusch (2020)).

It is important to point out that our documented local commonality of interest-rate movements is not inconsistent with a global factor structure per se. In fact, it is an intuitive observation that is consistent with the implications of most commonly specified global factor structures. Therefore, local commonality is not evidence against a global factor structure, but rather a more universally robust observation that does not depend on the exact specification of any particular global factor structures. By relying more on the universal robustness of the local commonality observation and less on the accuracy of a particular global factor structure specification, we seek to achieve more robustness in butterfly construction and investment performance.

4.2 Constructing Decentralized Butterflies with Nearby Maturities

With local commonality in mind, we propose to build decentralized butterfly portfolios with bonds of nearby maturities. Regardless of the underlying dynamics, since interest rates at nearby maturities show stronger co-movements, butterflies formed with nearby-maturity bonds are likely to become more stable, with only a small proportion of idiosyncratic movements left. Butterflies constructed with nearby maturities are also less intertwined with other butterflies and are therefore more decentralized, allowing larger diversification benefits when investing in multiple butterflies across different reference maturities.

To construct a decentralized 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 any factor dynamics assumptions, but start by directly estimating the covariance matrix Σ_t of the changes in the three chosen interest-rate series in the portfolio. Then, we choose the weights on the three interest-rate series $\mathbf{b}_t$ to minimize the variance of the interest-rate portfolio,

$$\min_{\mathbf{b}_t} \mathbf{b}_t^\top \Sigma_t \mathbf{b}_t, \quad \text{subject to} \quad \mathbf{b}_t^\top \mathbf{i}_2 = 1, \tag{27}$$

where $\mathbf{i}_2$ is an indicator vector with all elements being zero except the second element being one, as a way of normalizing the weight of the center maturity to one. The solution is

$$\mathbf{b}_t = \Sigma_t^{-1} \mathbf{i}_2 / (\mathbf{i}_2^\top \Sigma_t^{-1} \mathbf{i}_2). \tag{28}$$

At each point in time t, one can use a rolling window to estimate the covariance matrix Σ_t of the three swap rate series and solve the weights according to (28). Alternatively, one can directly regress rate changes at the middle maturity against the two rate changes on the wings over the same rolling period. The slope estimates would be for $(-b_{t,1}, -b_{t,3})$, i.e., the negative of the first and third elements of the weight vector $\mathbf{b}_t$. The variance of the regression residuals would reflect the variance left in the butterfly interest rate portfolio. The closer the three maturities are to each other, the more highly correlated the three interest-rate series tend to be, and the smaller would be the residual variation left.

Each bond responds to its underlying yield movement via its duration exposure. To convert the interest-rate portfolio weight $\mathbf{b}_t$ into the weight of the corresponding bond $\mathbf{w}_t$, we adjust for the duration difference for the three bonds,

$$w_{t,i} = b_{t,i} \mathcal{D}_2 / \mathcal{D}_i, \quad \text{for} \quad i = 1, 2, 3, \tag{29}$$

where we normalize the weight of the bond at center maturity to one ($w_{t,2} = 1$) and convert the wing weights on interest rates to weights on the corresponding bonds based on their duration ratio to the center maturity.

Butterfly portfolios constructed with the above weights have minimum variation left. Recall the bond return decomposition in (7), where the carry term y_t is known ex ante, and the convexity term $c_t = \frac{1}{2} \mathcal{C} (dy_t)^2$ also becomes deterministic under diffusive moves as the term $(dy_t)^2$ converges to the variance rate $\sigma_t^2 dt$ 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 with portfolio weight vector $\mathbf{w}_t$, the return attribution for the butterfly portfolio can be written as,

$$\mathbf{w}_t^\top \frac{d\mathbf{B}_t}{\mathbf{B}_t} = \mathbf{w}_t^\top \mathbf{y}_t dt + \frac{1}{2}\mathbf{w}_t^\top \mathbf{c}_t - \mathcal{D}_2 \mathbf{b}_t^\top d\mathbf{y}_t, \tag{30}$$

where in the last term we replace the bond portfolio weight vector $\mathbf{w}_t$ with the corresponding interest-rate portfolio weight vector $\mathbf{b}_t$ via the relation in (29). Since the interest-rate portfolio weight $\mathbf{b}_t$ is chosen in (27) to minimize the variance rate of the yield combination $\mathbf{b}_t^\top d\mathbf{y}_t$, the thus-constructed bond portfolio also has its expected return variance minimized.

4.3 Maturity Gaps and Stability of Butterfly Portfolios

To examine the effectiveness of butterfly construction in removing systematic interest rate risk, we treat the swap rates as the coupon rates or par yields of the corresponding par bonds and estimate the butterfly weights $\mathbf{b}$ on different swap rate combinations based on full sample regression on weekly changes of the associated swap rates. The regression residual represents the residual risk of the butterfly swap rate portfolio $\mathbf{b}^\top d\mathbf{y}_t$. At each reference maturity, we construct butterfly interest rate portfolios 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.

Table II reports in Panel A the annualized volatility of the weekly changes of the reference swap rates and the butterfly swap rate portfolios constructed around the reference maturity with increasing maturity gaps, all in percentage points. We have 9 swap rate series from 2 to 30 years, with 7-year as the median maturity. Thus, when the reference maturity is 7 year, the wings of the butterfly can be constructed with maturity gaps from 1 to 4. At 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 1 to 4.

[Table II about here.]

Weekly changes in the reference swap rate series have annualized volatility estimates from 0.866% to 0.934%. When we construct butterfly swap rate portfolios with adjacent maturities, the variance reduction is drastic. The volatility estimates of the weekly changes of the butterfly swap rate portfolios range from 0.064% to 0.109%. The variance ratios of the butterflies to the reference swap rates are merely 0.5-1.4%. Therefore, the butterfly construction with adjacent maturities removes 98.6-99.5% of the swap rate change variance, resulting in remarkably stable interest-rate portfolios.

As the maturity gap increases, the volatility estimates for the butterfly portfolios increase, a sign of deterioration in the effectiveness of variance reduction. When the maturity choice becomes the widest apart at the 7-year reference maturity with a maturity gap of 4, the volatility estimate reaches 0.242%. The variance ratio to the reference swap rate increases to 6.7%.

Another way of examining the stability of the butterfly interest rate portfolios is to examine their mean-reversion behavior. Panel B of Table II compares the autocorrelation of the weekly changes of the reference swap rates to the autocorrelation of weekly changes of the butterfly interest-rate portfolios. The autocorrelation estimates for weekly changes of the reference swap rates are close to zero, reflecting the high persistence of the swap rate series and the hard-to-predict nature of their weekly changes. By contrast, the autocorrelation estimates for the weekly changes of butterfly interest-rate portfolios constructed with adjacent maturities become much more negative, ranging from -39.5% to -52.4%. A well-constructed butterfly interest-rate portfolio can remove the most persistent component of the interest-rate movements and leave a portfolio that exhibits much stronger mean reversion. Butterfly interest rate portfolio construction with adjacent maturities is very effective in removing the persistent movements.

As the maturity gap increases, the effectiveness declines. When the reference maturity is at 7-year and the butterfly is constructed with the widest apart maturities, the autocorrelation estimate for weekly changes of the butterfly becomes very close to zero at -0.029, not much different from the autocorrelation estimate for the weekly changes of the reference swap rate series at -0.025. Therefore, when the maturity gap is wide apart, the butterfly construction may completely lose its effectiveness in removing the persistent component of the swap rate movement.

The statistics in Table II 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 99% of the variance in the reference swap rate movements. The construction also effectively removes the most persistent component of the swap-rate movements, resulting in an interest-rate portfolio that is much more mean-reverting than any single interest-rate series in the portfolio. This high mean-reverting behavior dictates that the value of the butterfly portfolio shows long-run stability as the variance of the butterfly movements does not increase with the holding horizon as fast as for the individual swap rate series. Therefore, when there are opposite demands at nearby maturities, broker dealers often use butterfly construction 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 by timing the strong mean-reverting behavior of the butterflies.

The analysis also shows the crucial importance of shrewd maturity choice. While butterflies constructed with nearby maturities show remarkable stability, butterflies constructed with far-apart maturities can be as unpredictable as the individual interest-rate series. The importance of maturity choice is an aspect often neglected in traditional model-centric analysis. If the yield curve variation can be well-captured by a two-factor structure, the model-centric thinking goes, one can hedge away the variation of the two factors with the combination of *any* three distinct yields. In this mindset, model design is of top importance whereas maturity choice is irrelevant. When a two-factor structure looks insufficient, the researcher will attempt a three-factor or an even higher dimensional structure, holding on to the mathematical trusim that K factors can always be neutralized with the combination of *any* $(K+1)$ yields, rendering maturity choice irrelevant again. The issue with such an approach is that the number of common factors itself depends crucially on the maturity choice for the analysis. This dependence makes any assumed global factor structures decidedly local to the data set used for the analysis. Furthermore, as the number of factors becomes large, the portfolios constructed with increasingly many legs lose the practical simplicity and flexibility of butterfly portfolios for liquidity provision.

The decentralized nature of our new theory prompts us to think from a different perspective, and allows us to pay more attention to maturity choice than dynamics assumptions. Our analysis shows that as long as one chooses nearby maturities, butterfly construction can be very effective in removing persistent interest-rate risk, regardless of what the true underlying factor structure is. Indeed, by relying on the local commonality of payoff structures, our new decentralized theory prompts us to ignore the global factor structure all together and construct stable butterflies that are independent of any structural model assumptions. It also allows us to construct each butterfly on its own, decentralized from the construction of other butterflies.

4.4 Mean-variance Investments in Decentralized Butterflies

The previous subsection highlights the importance of maturity choice in forming stable butterfly portfolios and shows that decentralized butterflies constructed with nearby maturities are naturally stable due to the local commonality of their payoff structures. This subsection uses an out-of-sample investment exercise to illustrate how to invest in these stable, decentralized butterflies based on classic mean-variance analysis of the portfolio's excess returns.

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 $\widehat{\sigma}^2_{t,i}$ and compute the annualized convexity contribution $\widehat{a}_{t,i} = \frac{1}{2}\mathcal{C}_i\widehat{\sigma}^2_{t,i}$ for each series i. 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 wings of the butterfly. The regression slope estimates represent estimates for $(-\widehat{b}_{t,1}, -\widehat{b}_{t,3})$. From the regression residual series, we obtain an annualized variance rate estimator $\widehat{\sigma}^2_{t,e}$ for the butterfly interest-rate portfolio $e_t = \mathbf{b}_t^\top d\mathbf{y}_t$.

According to the bond return attribution in (30), the variance of the returns on the par bond butterfly portfolio is related to the variance of the interest-rate portfolio by

$$\widehat{v}_{t,f} = \mathcal{D}_2^2\widehat{\sigma}^2_{t,e}. \tag{31}$$

With the weight on the reference bond at the center maturity normalized to one, the bond portfolio return variance is proportional to the residual variance of the interest-rate portfolio $\widehat{\sigma}^2_{t,e}$, scaled by the squared duration of the reference bond $\mathcal{D}_2^2$.

To determine the annualized expected excess return on the butterfly bond portfolio, we take expectation on the bond return decomposition equation in (7), annualize the expected return, and deduct the financing rate from it,

$$EER_{t,f} = \mathbf{w}_t^\top(\mathbf{y}_t - r_t) + \mathbf{w}_t^\top \widehat{\mathbf{a}}_t - \mathcal{D}_2 \mu_{t,e}, \tag{32}$$

where the first term captures the carry contribution over the financing rate, which we observe, and the second term captures the contribution from convexity, which we construct based on the variance estimators. Since the expected rate of change on each individual interest-rate series $\mu_t = \mathbb{E}_t[dy_t]/dt$ is inherently difficult to predict, we avoid predicting the expected rate of change on each individual series, but exploit the high mean-reverting behavior of the butterfly interest-rate portfolio and directly predict the expected rate of change on the interest-rate portfolio, $\mu_{t,e} = \mathbb{E}_t\left[\mathbf{b}_t^\top d\mathbf{y}_t\right]/dt$. For this purpose, 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, $\widehat{\mu}_{t,e} = -52e_t$.

Given the conditional mean excess return estimate in (32) and the conditional variance estimate in (31), we perform classic mean-variance investing by setting the investment weight on the butterfly par bond portfolio proportional to the mean-variance ratio,

$$n_{t,f} = s\widehat{EER}_{t,f}/\widehat{v}_{t,f}. \tag{33}$$

where s denotes a constant scaling coefficient. We apply the same scaling coefficient to different butterfly portfolios. The choice of the scaling coefficient is immaterial to our analysis. We choose a level so that the variations of the weights are within the usual confines of the leverage levels employed by institutional fixed income investors.

Table III reports the summary statistics of the allocation weights $n_{t,f}$ to butterfly par bond portfolios constructed with adjacent maturities around each of the six reference maturities. The statistics include percentile values of the allocation weights at 10, 25, 50, 75, and 90 percentiles.

The median weights are close to zero and the percentile values reveal a reasonably symmetric distribution. The variations of the allocation weights are larger at shorter maturities than at longer maturities.

[Table III about here.]

To examine how the allocation weights to different butterflies co-move, the last row of Table III reports the average cross-correlation estimates of each allocation weight series with the other allocation weight series. The average cross-correlation estimates are close to zero, suggesting that the variations of the estimated expected excess returns from these decentralized butterfly portfolios are largely independent of one another. This behavior forms a sharp contrast with the extremely high cross-correlation between the market price of risk estimates extracted from different par bonds. The butterfly construction has largely removed the common movements in the par bonds, so much so that even though the expected excess returns on different par bonds tend to move strongly together, the expected excess returns on the decentralized butterfly par bond portfolios are largely independent of one another.

Table IV reports the summary statistics of the weekly excess returns from investing in the decentralized butterfly par bond portfolios constructed with adjacent maturities. The investment performances are similar across butterflies with different reference maturities. The annualized mean excess returns range from 7% to 14%. The annualized standard deviations range from 9% to 12%. The annualized information ratios, defined as the ratio of annualized mean excess return to the annualized standard deviation, range from 0.84 to 1.18. The excess returns show large excess kurtosis and positive skewness.

[Table IV about here.]

The last row of Table IV reports the average cross-correlation estimates of each excess return series with other excess return series. Just as the allocation weights across different butterflies are largely independent of one another, so are the realized excess returns from the different butterfly portfolios. The near-independent nature of the decentralized butterflies offers strong diversification benefits when one simultaneously invests in multiple butterflies at different reference maturities. The last column reports the excess return statistics from an equal-weighted portfolio of the six butterflies across the six reference maturities. While the mean excess return retains the average magnitude of 10%, the standard deviation of the aggregate portfolio becomes much smaller than that of the individual butterflies at merely 5%. The annualized information ratio reaches 2.14, nearly doubling that of the individual butterfly investment.

We repeat the investment exercise on butterflies constructed with wider maturity gaps. Table V reports the summary statistics of the excess returns from investing in butterflies constructed with maximum maturity gaps from two in Panel A to four in Panel C. Compared to butterflies constructed with adjacent maturities in Table IV, the annualized information ratios of the investments on butterflies with wider maturity gaps become significantly lower, and increasingly so as the maturity gap becomes wider. The performance deterioration comes from both lowered mean excess return and increased return volatility. With larger maturity gaps, the excess return distribution also shows less positive or even negative skewness.

[Table V about here.]

Worse yet, as the maturity gap becomes wider, butterflies with different reference maturities become intertwined on the wings and lose their decentralized nature. As a result, the investment returns become more correlated with one another. The average cross-correlation among the excess return series is merely 9% when the maturity gap is 1, but increases to 23% when the maximum maturity gap is 2, to 37% when the maximum maturity gap is 3, and reaches as high as 47% when the maximum maturity gap is 4. The increased correlation reduces the benefit of diversification when one invests in multiple butterfly portfolios across different reference maturities. With the combined effects of deteriorated individual butterfly investment performance and reduced diversification benefits across different butterfly portfolios, the annualized information ratio for the equal-weighted butterfly portfolio declines quickly as the maturity gap increases.

4.5 Statistical Arbitrage Investment with Dynamic Term Structure Models

For comparison, we also perform an analogous out-of-sample investment analysis on butterflies constructed to benefit from the statistical arbitrage opportunities identified from estimating a dynamic term structure model. We estimate both a two-factor model as in Duarte, Longstaff, and Yu (2007) and a three-factor model as in Bali, Heidari, and Wu (2009). With the two-factor model, we construct butterflies to neutralize both interest-rate factors. With the three-factor model, we construct butterflies to neutralize the two more persistent factors. We adopt the parsimonious dimension-invariant cascade term structure specification of Calvet, Fisher, and Wu (2018), and follow their procedure in estimating the model parameters with quasi-maximum likelihood and extracting the factors with unscented Kalman filter. The parsimony of the specification enhances the parameter identification. To achieve better pricing performance, we re-estimate the model parameters once a year using data from the previous year.

The out-of-sample exercise follows a similar procedure as in the previous subsection. Starting on January 3rd, 1996, at each Wednesday t, we take the model parameter estimates from the previous year, run the unscented Kalman filter through the history up to date t to generate the history of the interest-rate factors. Based on the factor values at time t, we compute the sensitivity of the swap rates to the two factors and compute the butterfly portfolio weight $\mathbf{b}_t$ to neutralize the portfolio's exposure to these two factors while normalizing the weight of the center maturity to one. We then convert the swap rate portfolio weight to the corresponding par bond portfolio weight by adjusting for the par bond duration as in (29) .

We determine the allocation weight to a butterfly at each date to benefit from the statistical arbitrage opportunities as manifested by the pricing errors. We measure the difference between the observed value of the butterfly portfolio of swap rates and the corresponding model value, and treat the pricing error as temporary market mispricing. The more positive the pricing error, the higher the coupon payments of the par bond portfolio relative to the fair model value. To mitigate the impact of persistent mispricing due to model misspecification, we demean the pricing error by its historical average over the past year, and we scale the demeaned pricing error by the historical variance estimator of the pricing error series over the past year, multiplied by the squared duration of the reference swap contract.

The reference maturity choice, the maturity combination, and the excess return calculation all follow the same procedure as in the previous subsection. Table VI summarizes the out-of-sample annualized information ratio for the statistical arbitrage investment in each butterfly portfolio. When the butterflies are constructed with adjacent maturities, the statistical arbitrage trading generates reasonable out-of-sample performance. The information ratio for the aggregate portfolio reaches 1.54 when the butterflies are constructed with the two-factor model, and 1.84 when they are constructed with the three-factor model.

[Table VI about here.]

Even with a centralized dynamic term structure model, the investment performance remains highly dependent on maturity choice. Table VI shows that as the maturity gap of the butterflies becomes wider, the investment performance deteriorates quickly. Indeed, when the butterfly is constructed around the 7-year maturity with the widest maturity gap using the two-factor model, the average excess return from the out-of-sample investment exercise becomes negative. Model centralization does not replace the need for careful maturity consideration.

In a potentially high or even infinite dimensional world, the ambitious effort to build a robust and universally applicable global factor structure is likely futile. The elusive nature of a global factor structure also puts in question the true nature of the "arbitrage opportunities" identified from a no-arbitrage dynamic term structure model. Unless under certain special circumstances when a market maker can lock in a nonnegative future payoff without a cost, "arbitrage" opportunities identified from any assumed dynamic term structure models are most likely just a mirage. Investing in them as if they were truly arbitrage opportunities can be dangerous and can lead to financial ruins. Instead, owning up to the reality that they are not truly arbitrage opportunities does not preclude us from investing in them simply as an investment opportunity with potentially attractive risk-return tradeoff characteristics. Carefully gauging their expected returns and estimating their risks can potentially lead to more honest assessments, more prudent investments, and long-run sustainability. Our decentralized theory provides such a framework for doing mean-variance analysis on the excess returns of bonds and bond portfolios.

5. Concluding Remarks

In this paper, we propose a new modeling framework that is particularly suited for performing decentralized risk-return analysis on a bond or bond portfolio. The framework does not try to model the full dynamics of a short rate, or determine the functional form of the whole term structure, or its macroeconomic determinants, but rather focuses squarely on the behavior of the yield of the particular bond in question, and attributes the short-term investment return on the bond to variations of the bond's yield. The theory links the pricing of the bond's yield at any given point in time to the market pricing of risk on that bond and the conditional mean and variance forecasts on the bond's yield change at that time.

We illustrate the application of the theory in a decentralized investment analysis of a single par bond. The theory allows us to extract the market price of the par bond at any given point in time based on the observed yield spread over the financing cost at that time and the conditional mean and variance forecasts on the bond's yield change. We construct a simple conditional mean yield change forecast on the par bond based on the yield's sliding effect on the yield curve by assuming that changes in long-term floating interest rates are difficult to predict. We construct the conditional variance forecast with a simple historical variance estimator. The risk premium estimator based on these simple conditional forecasts generates better out-of-sample predicting performance on long-term bond excess returns than do commonly specified predictive regressions using the yield curve information.

We also apply the theory to the construction and investment of decentralized butterfly bond portfolios. The decentralized nature of our theory prompts us to pay more attention to maturity choice than dynamics assumptions in the butterfly construction. We propose the concept of local commonality and show that decentralized butterfly portfolios constructed with nearby maturities are much more stable than those constructed with maturities far apart. Global factor structure specification does not replace the need for shrewd maturity choice. With shrewd maturity choice, we can build stable decentralized butterfly bond portfolios and make profitable investments in them based on classic mean-variance analysis, without specifying any global factor structures.

The decentralized theory moves away from cross-sectional relative valuation, and puts renewed emphasis on classic risk-return analysis. While we retain the idea of forming portfolios of securities with related payoffs to cancel out a large proportion of their common movements, we do not put so much trust on any given no-arbitrage model to the point of regarding the pricing errors of the model as truly arbitrage opportunities. Instead, we propose to rely less on factor model construction, but pay more attention to security selection based on their structural linkages and our domain expertise, construct the portfolios based on their statistical behaviors and our insights about their near-term behaviors, and make investment decisions based on classic risk-return analysis.

Data Availability Statement

The empirical analysis in this paper uses US dollar LIBOR and swap rates, which can be downloaded from subscribed Bloomberg terminals.

Supplementary Material

An appendix is available at Review of Finance online.

References

Balduzzi, P., Das, S., Foresi, S., 1998. The central tendency: A second factor in bond yields. *Review of Economics and Statistics* 80, 62–72.

Bali, T., Heidari, M., Wu, L., 2009. Predictability of interest rates and interest-rate portfolios. *Journal of Business and Economic Statistics* 27, 517–527.

Bauer, M. D., Rudebusch, G. D., 2020. Interest rates under falling stars. *American Economic Review* 110, 1316–1354.

Bianchi, D., Büchner, M., Tamoni, A., 2021. Bond risk premiums with machine learning. *Review of Financial Studies* 34, 1046–1089.

Black, F., 1995. Interest rates as options. *Journal of Finance* 50, 1371–1376.

Calvet, L. E., Fisher, A. J., Wu, L., 2018. Staying on top of the curve: A cascade model of term structure dynamics. *Journal of Financial and Quantitative Analysis* 53, 937–963.

Campbell, J. Y., Shiller, R., 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* 58, 494–514.

Chance, D. M., Jordan, J. V., 1996. Duration, convexity, and time as components of bond returns. *Journal of Fixed Income* 6, 88–96.

Christensen, P. O., Sorensen, B. G., 1994. Duration, convexity and time value. *Journal of Portfolio Management* 20, 51–60.

Cieslak, A., Povala, P., 2015. Expected returns in treasury bonds. *Review of Financial Studies* 28, 2859–2901.

Clark, T. E., McCracken, M. W., 2012. Advances in forecast evaluation. In: *Handbook of Economic Forecasting*, ed. by Timmermann, A., Elliott, G. Elsevier, Amsterdam.

Cochrane, J. H., Piazzesi, M., 2005. Bond risk premia. *American Economic Review* 95, 138–160.

Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. *Review of Financial Studies* 22, 2801–2833.

Cox, J. C., Ingersoll, J. E., Ross, S. R., 1985. A theory of the term structure of interest rates. *Econometrica* 53, 385–408.

Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.

Duarte, J., Longstaff, F. A., Yu, F., 2007. Risk and return in fixed income arbitrage: Nickels in front of a streamroller. *Review of Financial Studies* 20, 769–811.

Duffee, G. R., 2011. Information in (and not in) the term structure. *Review of Financial Studies* 24, 2895–2934.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. *Mathematical Finance* 6, 379–406.

Dybvig, P. H., Ingersoll, J. E., Ross, S. A., 1996. Long forward and zero-coupon rates can never fall. *Journal of Business* 69, 1–25.

Fama, E. F., Bliss, R. R., 1987. The information in long-maturity forward rates. *American Economic Review* 77, 680–692.

Gargano, A., Pettenuzzo, D., Timmermann, A., 2017. Bond return predictability: Economic value and links to the macroeconomy. *Management Science* 65, 508–540.

Giglio, S., Kelly, B. T., 2018. Excess volatility: Beyond discount rates. *Quarterly Journal of Economics* 133, 71–127.

Greenwood, R., Vayanos, D., 2014. Bond supply and excess bond returns. *Review of Financial Studies* 27, 663–713.

Harvey, D. I., Leybourne, S. J., Newbold, P., 1997. Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13, 281–291.

Heidari, M., Wu, L., 2010. Market anticipation of federal reserve policy changes and the term structure of interest rates. *Review of Finance* 14, 313–342.

Hua, J., Wu, L., 2018. Monetary policy rule as a bridge: Predicting inflation without predictive regressions. *Journal of Financial and Quantitative Analysis* 53, 2559–2586.

Joslin, S., Priebsch, M., Singleton, K. J., 2014. Risk premiums in dynamic term structure models with unspanned macro risks. *Journal of Finance* 69, 1197–1233.

Kim, D. H., Orphanides, A., 2012. Term structure estimation with survey data on interest rate forecasts. *Journal of Financial and Quantitative Analysis* 47, 241–272.

Krippner, L., 2012. Measuring the stance of monetary policy in zero lower bound environments. *Economics Letters* 118, 135–138.

Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1, 54–61.

Ludvigson, S. C., Ng, S., 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22, 5027–5067.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.

Rapach, D. E., Zhou, G., 2013. Forecasting stock returns. In: *Handbook of Economic Forecasting*, ed. by Elliott, G., Timmermann, A. Elsevier, Amsterdam.

Vasicek, O. A., 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 177–188.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21, 1455–1508.

Woodford, M., 1999. Optimal monetary policy inertia. *The Manchester School* 67, 1–35.

Table I. Out-of-sample forecasting performance on par bond excess returns

Entries report the out-of-sample forecasting performance on six-month excess returns on par bonds with maturities from 2 to 30 years. Panel A reports the performance of the Cochrane-Piazzesi (CP) bond risk premium estimator based on a portfolio of forward rates. Panel B reports the performance of our bond risk premium estimators (CW) based on an implementation of our decentralized theory. The performance measures include both a forecasting R^2 measure and the Diebold–Mariano (DM) (1995) t-statistic. The risk premiums are estimated with a 10-year rolling window. The out-of-sample statistics for each series are computed on 704 weekly observations from January 5th, 2005 to Jun 27th, 2018.

	Panel A. CP		Panel B. CW	
Maturity	R^2	DM	R^2	DM
2	-0.35	-0.96	0.13	0.57
3	-0.43	-1.61	0.14	1.15
4	-0.41	-1.61	0.18	1.88
5	-0.39	-1.51	0.19	2.26
7	-0.31	-1.30	0.18	2.59
10	-0.21	-1.02	0.15	2.73
15	-0.13	-0.76	0.12	2.75
20	-0.08	-0.59	0.10	2.78
30	-0.06	-0.47	0.09	2.82

Table II. Effectiveness of variance reduction via butterfly construction

Entries report the annualized volatility in panel A and weekly autocorrelation in panel B of weekly changes on each reference swap rate series, as well as the butterfly swap rate portfolios constructed around the reference maturity with increasing maturity gaps. 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.

Reference	Reference	Butterflies with maximum maturity gap of			
Maturity	Swap	1	2	3	4
		Panel A. Annualized volatility of weekly changes			
3	0.876	0.078	0.089	0.106	0.122
5	0.928	0.079	0.115	0.192	0.213
7	0.934	0.078	0.123	0.169	0.242
10	0.928	0.109	0.147	0.191	0.213
15	0.893	0.064	0.108	0.124	0.135
20	0.866	0.070	0.102	0.107	0.113
		Panel B. Autocorrelation of weekly changes			
3	-0.001	-0.395	-0.326	-0.231	-0.179
5	-0.018	-0.524	-0.219	-0.073	-0.036
7	-0.025	-0.442	-0.247	-0.097	-0.029
10	-0.033	-0.412	-0.318	-0.243	-0.200
15	-0.050	-0.439	-0.316	-0.262	-0.232
20	-0.049	-0.443	-0.414	-0.401	-0.384

Table III. Summary statistics of time-varying allocation weights to butterfly par bond portfolios

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

Maturity	3	5	7	10	15	20
10	-42.89	-23.80	-24.92	-15.36	-19.15	-19.35
25	-16.51	-9.29	-7.97	-5.14	-5.51	-5.38
50	-0.43	-0.15	-0.25	0.05	-0.05	0.00
75	14.94	7.24	7.55	4.67	6.73	5.29
90	40.44	21.07	23.21	15.31	24.57	18.98
Corr	-0.04	0.01	0.07	0.09	-0.05	-0.06

Table IV. Investment excess returns on butterfly par bond portfolios with adjacent maturities

Entries report the summary statistics of excess returns from investments in butterfly par bond 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 across the six reference maturities, and the correlation value in the last column represents the grand average of the cross-correlations among the six butterfly investment return series.

Maturity	3	5	7	10	15	20	Aggregate
Mean	0.11	0.11	0.07	0.08	0.09	0.14	0.10
Stdev	0.10	0.09	0.09	0.09	0.09	0.12	0.05
IR	1.11	1.18	0.84	0.92	1.02	1.14	2.14
Skewness	2.53	1.15	9.47	0.93	1.55	1.03	0.78
Kurtosis	29.95	31.87	189.46	45.97	62.31	25.71	25.11
Corr	0.02	0.09	0.11	0.12	0.08	0.09	0.09

Table V. Investment excess returns on butterfly bond portfolios with wider maturity gaps

Entries report the summary statistics of excess returns from investments oin butterfly bond portfolios constructed with wider maturity gaps, with each panel representing butterflies with a different maximum maturity gap. 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 return series.

Maturity	3	5	7	10	15	20	Aggregate
Panel A. Maximum maturity gap = 2							
Mean	0.10	0.08	0.05	0.05	0.08	0.11	0.08
Stdev	0.11	0.11	0.10	0.10	0.12	0.14	0.07
IR	0.87	0.78	0.53	0.54	0.65	0.79	1.18
Skewness	2.02	0.96	-1.70	-4.07	-0.30	0.39	-2.00
Kurtosis	22.53	20.94	35.02	80.67	35.20	28.16	38.94
Corr	0.01	0.13	0.34	0.32	0.35	0.23	0.23
Panel B. Maximum maturity gap = 3							
Mean	0.09	0.07	0.05	0.06	0.06	0.11	0.07
Stdev	0.14	0.15	0.11	0.13	0.12	0.14	0.09
IR	0.68	0.46	0.42	0.44	0.51	0.81	0.82
Skewness	1.24	1.04	-3.51	-4.25	-1.30	-0.17	-2.74
Kurtosis	18.52	23.20	63.53	85.68	49.94	32.65	47.18
Corr	0.11	0.32	0.48	0.49	0.46	0.37	0.37
Panel C. Maximum maturity gap = 4							
Mean	0.09	0.06	0.05	0.06	0.06	0.10	0.07
Stdev	0.14	0.15	0.14	0.13	0.13	0.14	0.10
IR	0.62	0.40	0.33	0.45	0.45	0.72	0.66
Skewness	0.98	0.86	-1.86	-4.45	-2.25	-0.14	-2.47
Kurtosis	13.80	24.55	37.24	88.69	56.92	34.10	43.15
Corr	0.22	0.48	0.61	0.57	0.52	0.43	0.47

Table VI. Statistical arbitrage investment in butterfly bond portfolios with term structure models

Entries report the annualized information ratio for investing in butterfly par bond portfolios to benefit from statistical arbitrage opportunities identified from the pricing errors of a two-factor (Panel A) and a three-factor (Panel B) dynamic term structure model, respectively. Each column denotes a butterfly with a different reference maturity. Each row denotes the maximum maturity gap the butterfly is constructed around the reference maturity. The last column reports the information ratio on the aggregate portfolio constructed with equal weighting on the six butterfly portfolios.

Reference:	3	5	7	10	15	20	Aggregate
Maturity gap:							
	Panel A. Butterflies constructed with a two-factor model						
1	0.65	1.12	0.86	0.57	0.83	0.74	1.54
2	0.66	0.52	0.29	0.33	0.51	0.74	0.86
3	0.50	0.16	0.16	0.29	0.39	0.81	0.56
4	0.40	0.07	-0.02	0.24	0.34	0.66	0.35
	Panel B. Butterflies constructed with a three-factor model						
1	0.59	1.25	0.70	1.12	0.90	0.84	1.84
2	0.60	0.67	0.34	0.78	0.56	0.79	1.10
3	0.50	0.08	0.14	0.51	0.63	0.66	0.68
4	0.46	0.06	0.19	0.45	0.59	0.66	0.60

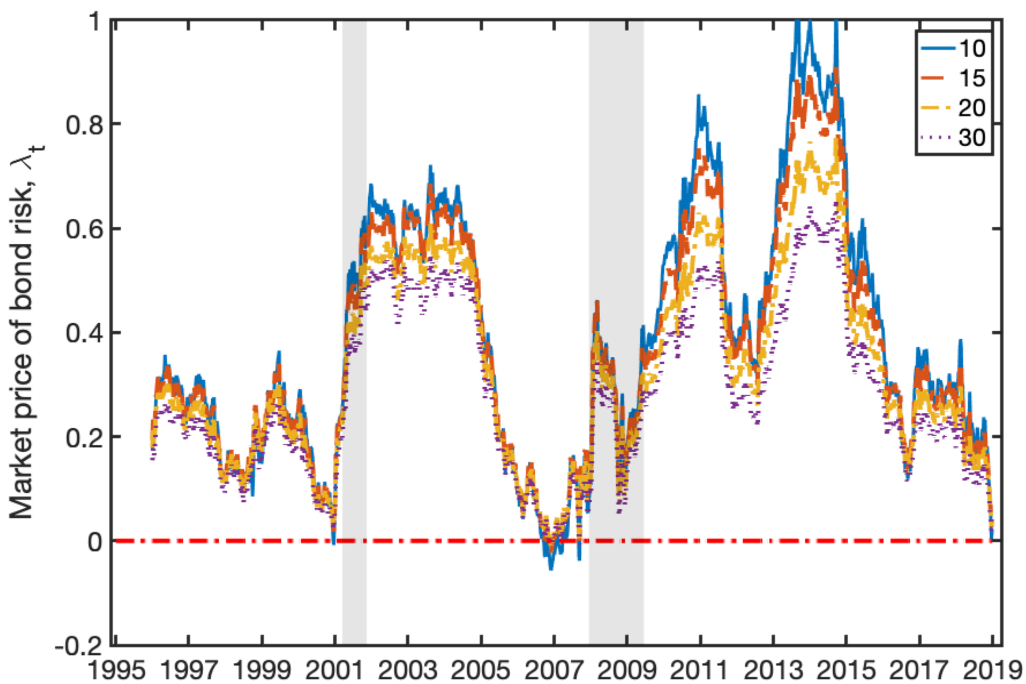

Figure 1. Market price of bond risk extracted from long swap rates based on an implementation of the decentralized theory. The swap rate data are weekly from January 1995 to December 2018. The first year of the data are used to estimate the rolling-window variance rate of swap rate changes.

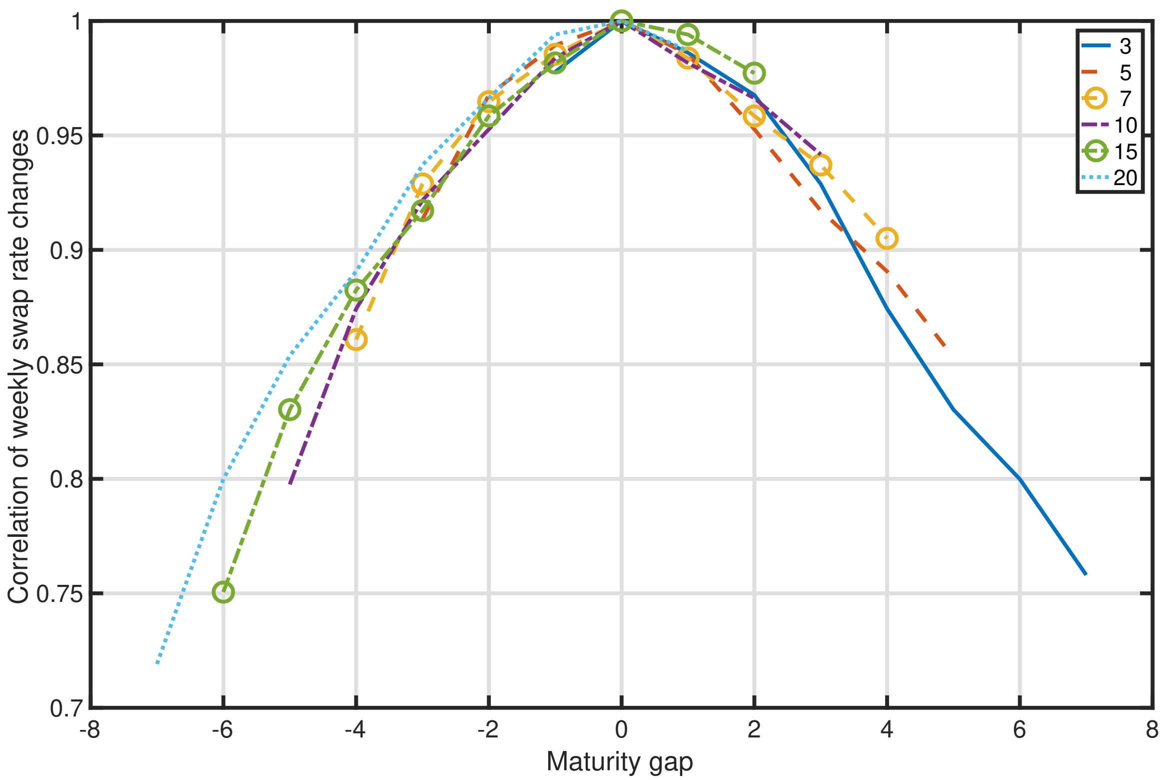

Figure 2. Cross-correlations of weekly changes of swap rates at different maturity gaps. The six lines denote correlation estimates of weekly swap rate changes between each of the six reference maturities at 3, 5, 7, 10, 15, and 20 years and other maturities. The correlations are estimated on the weekly data series from January 1995 to December 2018. The estimates are plotted against the maturity gap between the other maturities and the reference maturity.