Extracting Expectations in Affine Term Structure Models

Halil I. Aydin and Yildiray Yildirim *

Current version: August 22, 2015

Abstract

In this paper, we study the problem of implementation of [Ross (2015)] Recovery theorem to disentangle the pricing kernel and physical probabilities from observed bond yields within discrete time affine term structure models. As a remedy to the problem of obtaining Arrow-Debreu prices of state transitions, we propose Markov chain approximation to autoregressive processes. Our work suggests that affine setting offers rich structure that enables us to obtain necessary inputs in empirical applications. In the second part, we estimate a canonical discrete time Gaussian three factor term structure model with the U.S. Treasury bond yields. We decompose bond yields into expectation and risk components without specifying risk adjustment inside the model. The results indicate that power of term spread in predicting economic activity stems from level of expectations component and change in risk premium component.

*Aydin, halil.aydin@tcmb.gov.tr, Istanbul School of Central Banking, The Central Bank of Turkey; Yildirim, yildiray.yildirim@baruch.cuny.edu, Zicklin School of Business, Baruch College, CUNY, 137 E 22nd, New York 10010, USA. We thank the seminar participants at Whitman School of Business and Istanbul School of Central Banking for their helpful comments and suggestions.
1. Introduction

In this paper, we attempt to decompose long term bond yields into average market expectations and risk premium components according to the Recovery Theorem outlined by Ross (2015). We first revisit Ross (2015) approach and summarize the challenges one faces during the implementation. Then, we propose a new method to overcome these challenges within affine term structure models. We show that if the risk neutral evolution of the factors governing term structure is stationary, one can obtain Arrow-Debreu prices and hence disentangle expectations and risk premium.

Contrary to common tradition in asset pricing, Ross (2015) shows that, with minor assumptions, it is possible to recover the representative agent’s pricing kernel and beliefs about future asset returns from asset prices alone. As such, Ross (2015) suggests a new tool to reach physical dynamics despite the fact that pricing is carried out in risk neutral world. This procedure is the opposite approach of the traditional asset pricing models that require a specification of risk adjustment in changing pricing mechanism from risk neutral to physical. With the Recovery Theorem, we no longer need to specify the market price of risk to obtain physical probabilities from asset prices.

The Recovery Theorem is a significant breakthrough in asset pricing literature since market expectations and risk premium are invaluable source of information for policy makers as well as investors. With market assessment of future price movements, policy makers are able to monitor market developments in a timely manner. At the same time, investors are able to construct portfolios in accordance with market beliefs and to hedge against systematic risk. The Recovery Theorem enables us to utilize this rich source of information on market expectations and risk premium embedded in asset prices by reverse engineering of pricing components from observed market prices.

Furthermore, we explore the implications of the Recovery Theorem in a completely new setting. Specifically, we use affine term structure models because they offer first hand solutions to problems we face during implementation. In this respect, this paper is the first attempt of application of the Recovery Theorem to affine term structure models. We show that using Markov chain approximation to risk neutral processes, we are able to recover physical dynamics in affine term structure models.

More specifically, we pose two research questions. The first question is “How can we apply the Recovery Theorem in term structure models”? To answer this question we obtain Ross (2015)’s results in a discrete time model and discuss the challenges we face during
implementation. Depending on the answer of the first question, we ask the second question: “What are the empirical findings if the Recovery Theorem is applied to term structure of bond yields”? In search for an answer to this question, we find that our approach enables us to decompose bond yields as expectations and risk premium components, where the expectation component has predictive content for economic activity.

Despite its noted benefits, the Recovery Theorem poses challenges in its practical implementation because contingent prices for all states of the world are not observable. In order to implement the theorem, we need first to decide on the possible states of the world and then form the Arrow-Debreu price matrix that maps transitions between these states. This requirement is our starting point and our solution to this problem forms the main contribution of this paper.

In order to identify the challenges that the Recovery Theorem brings in estimation, we revisit Ross (2015)’s modeling assumptions. The first assumption is the existence of a finite state Markov chain driving the states of the world such that the probabilities attached to movements between those states are known and represented by a transition matrix. Second, uncertainty surrounding future consumption is characterized by the states and the representative agent with time-separable additive utility function maximizes his intertemporal utility in this environment. Last but not least, in equilibrium, the ratio of marginal utilities of the representative agent across two dates is assumed independent of initial states. With these assumptions, equilibrium conditions are expressed in matrix form and the extraction of physical probabilities from discounted risk neutral probabilities (i.e., Arrow-Debreu prices) reduces to an eigenvalue problem. With the help of Perron-Frobenius theorem, Ross (2015) proves the existence of a unique solution to this eigenvalue problem.

The economic background and mathematical derivation of Ross recovery are well-founded, but we describe two issues in the implementation of recovery theorem. First, we need to obtain the matrix of Arrow-Debreu prices. To do so, we propose Markov chain approximation of stochastic processes governing risk factors of term structure of bond yields. Our approach is to utilize the information in yield curve to obtain Arrow-Debreu prices by approximating autoregressive processes as Markov chains. The procedure yields discrete state space and transition probabilities; altogether these constitute the Arrow-Debreu prices which we seek for implementation.

The second issue with the implementation of the Recovery Theorem is to have state-dependent discount factors. In fact, the difference between physical probabilities and risk neutral counterparts depends on this requirement. Ross (2015) shows that if the discount
factor is not state-dependent, the recovered probabilities are the same as the risk neutral ones. Hence, in order to differentiate physical and risk neutral probabilities, we need to allow state dependent discount factors. The affine structure enables us to model discount factors as one-period bond prices. We use this feature to convert risk neutral probabilities into Arrow-Debreu prices. This procedure yields state dependent-discount factors and (in the sense of the Recovery Theorem) has the desired properties under affine models. The resulting prices satisfy the second requirement and can be embedded into the Markov matrix since one-period bond prices depend on risk factors under affine setting. Thus, we show that both requirements of the Recovery Theorem are satisfied within affine setting and we are therefore able to apply the Recovery Theorem.

The benefits of using the Recovery Theorem become apparent in empirical applications. In the empirical part, we attack the problem of uncovering physical expectations using only the information about risk neutral dynamics. In contrast to the conventional way of reaching physical expectations through imposing the market price of risk, we are able to compute expectations from discounted risk neutral probabilities alone. This is possible because the Recovery Theorem requires only risk neutral probabilities and yields physical expectations as output. This unique feature of the Recovery Theorem makes it particularly attractive in empirical work since the price of risk specification has shown to affect physical expectations obtained in affine models.

In order to highlight the comparative advantage of the Recovery Theorem, we estimate a three-factor affine model. Then, we obtain physical expectations using bond yields of U.S. Treasury securities. This allows us to decompose long-term yields into the expectations of short rate and risk premium. This way, we obtain economically plausible expectation and risk premium estimates. We find interesting results regarding the evolution of the expectations and risk premium components. Our estimates produce a time-varying expectations component as opposed to a flat expectations component usually encountered in small samples with affine models. Also, we find changes in the risk premium to be significant in the GDP forecast regressions. Risk premium component shows volatile and time varying pattern indicating a counter cyclical behavior with business cycles as expected in the theoretical studies. These findings show the benefit of the Recovery Theorem and close some important gaps in the literature.

As an additional empirical exercise, we estimate another affine model to obtain inflation adjusted interest rate dynamics under risk neutral model using the cross section of U.S. Treasury Inflation Protected Securities. We are able to obtain physical expectations of inflation-indexed interest rate and inflation risk premium. Combined with nominal exercise,
we are able to decompose long-term yields into the real and inflation expectations and into the real and inflation risk premium. We show that inflation expectations obtained using this method produces substantial variation as opposed to alternative methods.

In sum, the contribution of our work has distinctive features. First and foremost, we obtain physical probabilities and risk adjustment using only the information available in the risk-neutral model. This approach relaxes the restrictions imposed by risk parameters and produces results independent of risk specifications. Second, the model enables risk adjustment to take place over large state spaces covering rare events and associated probabilities. Considering rare events provides additional robustness to model estimates. Third, our empirical results show additional benefits of using the Recovery Theorem in alleviating the econometric problems usually encountered in small samples. In contrast to previous attempts, our approach with Markov chains produces substantial variation in long-term expectations of bond yields, though the modeling assumptions remain minimal.

The remainder of this paper is organized as follows. We start in Section 2 with a review of the relevant literature. Section 3 gives theoretical background of Ross (2015) Recovery Theorem and explains the challenges with its implementation. Section 4 summarizes affine term structure models gives the details of the Markov chain approach. In Section 5, we apply the model empirically and report the findings. We discuss the implications of our empirical findings in Section 6. Finally, Section 7 concludes and offers directions for future research.

2. Literature

In affine models, one is able to decompose long-term bond yields into the expectation of short rate and risk premium. The origin of the decomposition is the deviation of the observed long term bond yields from expectations hypothesis\(^1\). The literature on the extraction of risk premium and expectations in affine setting documents several such attempts. Duffie and Kan (1996) and Dai and Singleton (2000) are the initial attempts of classifying multivariate interest rate models under the name “affine”. Duffee (2002) builds on these works and compares various affine models with different risk specifications. These models are used by researchers to fit observed bond yield curve with a small number of risk factors.

Early generation of affine models such as Dai and Singleton (2000) and Duffee (2002)

\(^1\)Cochrane (2001) documents three different definitions of expectation hypothesis leading to three different risk premia. Unless stated otherwise, we use the term risk premium to refer to bond yield premium in his notation.
work with latent (or statistical) risk factors. Supporting evidence for this approach are the findings of [Litterman and Scheinkman 1991]. They show that three factors such as the three largest principal components are sufficient to capture most of the variation in the yield curve. The three factors are interpreted as the level, slope and curvature factors due to their effects on the yield curve. Recently, [Cochrane and Piazzesi 2005] derive a new statistical factor from linear combination of forward rates and yields. They show that the new return-forecasting factor describes time variation in expected returns of all bonds and increases forecasting performance.

Later studies on affine models incorporate macroeconomic variables as additional risk factors besides the statistical factors. These hybrid models aim to increase the explanatory power of risk factors in the estimation of yield curve dynamics. [Ang and Piazzesi 2003] incorporate contemporary and lagged values of inflation and GDP growth rate as macroeconomic factors. They show that inclusion of these factors increases the predictive power of affine models.

Due to the rich structure of hybrid models with both statistical and macro factors affecting bond yields makes the estimation complicated. To make estimation tractable, [Ang and Piazzesi 2003] impose an additional restriction on the interaction between statistical and macro factors. Accordingly, the feedback effects from macro factors into statistical factors are excluded, implying that macro variables interact with each other but do not respond to changes in short and/or long interest rates. Hence, the expectations and risk premium components obtained from this model do not capture the effects of yield curve on macroeconomy.

[Bernanke, Reinhart, and Sack 2004] employ another affine model but assume that the state vector includes only macroeconomic factors, eliminating statistical factors. Both the short-rate expectations and the prices of risk are determined only by the macroeconomic factors. Since there is no statistical factors in the state dynamics, the empirical implementation of the model is simplified significantly. The framework captures effects of movements in the long-term yields driven by observable factors, but it does not empirically distinguish the role of long term yields from that of lagged macroeconomic variables. Also, similarly to [Ang and Piazzesi 2003], the model implies no effect of yield curve on the dynamics of the economy.

[Rudebusch, Sack, and Swanson 2007] and [Hördahl, Tristaní, and Vestin 2006] derive New-Keynesian models where monetary policy shocks, inflation target, and output gap are the macroeconomic factors affecting the yield curve dynamics. The state dynamics are
obtained from micro-foundations such as optimality conditions of households and firms. Thus the models have more theoretical basis and the interaction between factors affecting yield curve has economic foundations. However, the models bring substantial computational challenges due to lack of closed from solutions. Hence, the estimation of model with the actual data is challenging and risk premium obtained from these models does not show satisfactory performance.

Moreover, irrespective of the choice of factors whether purely statistical or macroeconomic, the estimation of affine models relies on the specification of risk adjustment (i.e., market price of risk) and identification of the transmission of shocks among factors. A common approach for the risk adjustment is to assume a linear pricing rule where the price of risk is proportional to underlying risk factors. Similarly, the identification of risk factors is achieved by unidirectional transmission of shocks from factors into bond yields, implying that risk factors have impact on bond yields whereas yields do not impact on the factors.

Overall, expected yield and risk premium depend on chosen risk factors and the identification of the relation between the factors and the yields. Duarte (2004), Duffee (2007) and Dai and Singleton (2002) emphasize that risk premium estimates in the literature depend on preferred model and structure that is imposed over risk factors. This is to say, the choice of risk factors as well as the relation between factors are important determinants of risk premium and hence expectation components. As a result, the findings in the literature are mixed and determination of the proper risk adjustment remains a challenge.

In order to disentangle the expectation and risk premium components in bond yields, the standard approach is to assume both expected yield and risk premium as affine functions of risk factors and to estimate both components jointly within the model. However, decomposition obtained from standard affine model is not satisfactory to be used in real time decisions. Kim and Orphanides (2012) elaborate on this challenge and attribute its causes to the complications arising in the estimation of physical parameters. The problem stems from the high persistence in risk factors and a high number of parameters to be estimated, and becomes even more severe when the models are estimated using small samples.

Kim and Orphanides (2012) extend the affine model by augmenting survey forecasts as an additional source of information in the estimation. Accordingly, survey forecasts are assumed to be noisy observations of market expectations. With this change, they are able to obtain more variation in the expectations components and to improve efficiency of the estimated parameters. But their approach is still prone to criticism raised by Duffee (2007) about model specific risk premium estimates.
The lack of statistical tools for uncovering physical parameters causes model estimates to attribute variations in long-term bond yields to changes in risk premium and to produce slowly changing expectations. This has important implications since survey expectations change more than do model implied expectations. Bauer, Rudebusch, and Wu (2014) show that slowly changing expectations do not contain any information for monetary policy about the future. In order to account for this effect, they suggest bias correction methods to increase the persistence of the factors. But this leads to unit root in the factors, allowing for negative interest rates. Additionally, the emphasis is only on the first factor. In our framework, we show that even with a persistence less than unity in the first factor we can reach economically plausible expectations components. In doing so, we do not need any bias correction since the Recovery Theorem allows us to obtain physical feedback matrix.

Our work improves upon the existing modes in the following important ways. We construct the link between affine models and the Ross’s Recovery approach. We select risk factors as traditional statistical factors but do not explicitly specify the market price of risk. This enables us to estimate risk factors and parameters simultaneously in the risk neutral setup. With the help of the Recovery Theorem, we find physical state dynamics yielding the expectations component. Finally, we infer risk premium as the difference between the observed yields and the average expectations of the short term interest rates.

3. Ross Recovery

The economic theory behind Ross (2015) starts with a two-period optimization problem for a representative agent. The economic uncertainty is characterized by a finite number of states of nature where each state corresponds to a different level of consumption. Accordingly, $c_{0i}$ denotes consumption at time $t = 0$ given that the state at time $t = 0$ is $X_0 = i$. Similarly, $c_{1j}$ is the consumption at time $t = 1$ given that state at time $t = 1$ realizes as $X_1 = j$. There are $m$ possible states of nature at each time period but the agent observes the state in which he is currently in. In other words, the uncertainty in the consumption exists for next period only.

The stochastic process governing the states of nature is assumed to be a Markov chain consisting of the true transition matrix indicating switching probabilities from state $X_0 = i$ to $X_1 = j$. The transition probabilities are called natural or physical probabilities and are denoted by $f_{ij}$. For each state realization in the next period, the agent gets one unit of consumption good and receives nothing in other states. These state-contingent claims (i.e.
Arrow-Debreu prices), $p_{ij}$, pay one unit of consumption good if the state $j$ is the actual realization at time $t = 1$. Lastly, there is no exogenous source of income other than the initial wealth, $w$, and $\beta$ denotes the subjective discount factor. With this information, the optimization problem in discrete time and discrete state space environment for a representative agent can be formulated as follows:

$$\max_{\{c_{0i}, c_{1j}\}} \left\{ U(c_{0i}) + \beta \sum_{j=1}^{m} U(c_{1j}) f_{ij} \right\}$$

(3.1)

subject to intertemporal budget constraint:

$$c_{0i} + \sum_{j=1}^{m} c_{1j} p_{ij} = w$$

(3.2)

The equilibrium solution to this problem can be found by constrained optimization assuming that requirements for an interior solution hold. The closed-from solution can be obtained by forming the Lagrangian which takes the form

$$\mathcal{L} \equiv U(c_{0i}) + \beta \sum_{j=1}^{m} U(c_{1j}) f_{ij} + \mu \left[ w - c_{0i} - \sum_{j=1}^{m} c_{1j} p_{ij} \right]$$

(3.3)

where $\mu$ denotes Lagrange multiplier. First order conditions, obtained from partial derivatives of the Lagrangian with respect to control variables $c_{0i}$ and $c_{1j}$ are given as

$$U'(c_{0i}) - \mu = 0$$

(3.4)

$$\beta U'(c_{1j}) f_{ij} - \mu p_{ij} = 0$$

(3.5)

The solution to this set of equations yields undiscounted pricing kernel, $\phi_{ij}$:

$$\phi_{ij} = \frac{p_{ij}}{f_{ij}} = \frac{\beta U'(c_{1j})}{U'(c_{0i})}$$

(3.6)

The equation (3.6) characterizes the evolution of consumption along the optimal path. It is the equilibrium condition in which a representative agent must be indifferent between consuming one more unit today and saving that unit and consuming in the future. The equilibrium condition holds for any two consecutive time periods $t$ and $t + 1$. [Ross (2015)] states that in a multi-period model with complete markets and state-independent, intertemporally additive, and separable preferences; there is a unique representative agent utility function satisfying the above optimum condition.
If we know the form of the utility function, the discount factor and the state-price densities, we would be able to find physical probabilities from kernel equation (3.6) as in Bliss and Panigirtzoglou (2004). However, physical probabilities obtained via this method depend on the chosen utility function.

Ross (2015) follows a different track and converts the pricing kernel equation into an eigenvalue problem. This allows him to find the discount factor and physical probabilities simultaneously. It is important to note that in this framework, there is no specific shape for the utility function such as constant absolute (and/or relative) risk aversion. The only requirement is to have state-independent, intertemporally additive separable preferences to be able to reach matrix representation in the optimization problem.

Ross (2015) builds on the solution (3.6) by adding one more assumption that the pricing kernel depends only on the marginal rate of substitution between future and current consumption (i.e. independent of initial state). With this additional assumption, the kernel takes the form

$$\phi_{ij} = \frac{P_{ij}}{f_{ij}} = \beta \frac{U'(c_j)}{U'(c_i)} \text{ for any } j = 1, 2, \ldots, m$$

(3.7)

This form of the kernel is the fundamental equation for the Recovery Theorem. Ross (2015) converts this equation into an eigenvalue problem to find physical probabilities solely from the matrix of Arrow-Debreu prices. Specifically, rewriting the kernel in a matrix form yields

$$DP = \beta FD$$

(3.8)

where $P$ is an $m \times m$ matrix of Arrow-Debreu prices, $F$ is an $m \times m$ matrix of physical transition probabilities, and $D$ is a diagonal matrix where diagonal entries are the inverse of the kernel $\phi_{ij}$ for $j = 1, 2, \ldots, m$. Since $F$ is a stochastic matrix, the elements of each row of matrix $F$ sum up to one. Using this information, Ross (2015) is able to rewrite (3.8) as an eigenvalue problem as follows

$$Pz = \beta z$$

(3.9)

where $z$ is an $m \times 1$ column vector such that $z = D^{-1}e$ and $e$ is an $m \times 1$ vector of ones.

Since matrix $P$ represents Arrow-Debreu prices, all elements of $P$ are nonnegative. Perron-Frobenius theorem states that for matrices with nonnegative elements, there is a unique positive eigenvalue $\beta$ and associated eigenvector $z$. Thus, if we have Arrow-Debreu prices covering all transitions between the states of the world, we are able to find subjective discount factor $\beta$ and the corresponding eigenvector $z$. Ultimately, physical probabilities $F$
can be obtained from these elements as follows

\[ F = \frac{1}{\beta} D P D^{-1} \quad \text{(3.10)} \]

The implementation of [Ross (2015)] Recovery Theorem requires knowing Arrow-Debreu prices covering all possible state transitions. This information is the key input during the implementation of [Ross (2015)] Recovery Theorem. Our approach to obtain this input is based on approximating the data generating process by Markov chains, as will be explained in the next section.

Before discussing the implementation of [Ross (2015)] approach, we rewrite Arrow-Debreu prices in terms of the risk neutral probabilities. This clarifies our approach outlined in the next sections. The price of a state contingent asset \( S_i \) with the next period price \( S_j \) can be written in terms of Arrow-Debreu securities as follows:

\[ S_i = \sum_{j=1}^{m} p_{ij} S_j \quad \text{(3.11)} \]

The equation follows from the Fundamental Theorem of Asset Pricing by [Dybvig and Ross (1987)]. When we insert the kernel into this equation, we get the price of asset \( S_j \) in terms of ratio of marginal utilities, physical transition probabilities, and the discount factor.

\[ S_i = \sum_{j=1}^{m} \beta \frac{U'(c_j)}{U'(c_i)} f_{ij} S_j \quad \text{(3.12)} \]

This equation, together with the analogue of the pricing equation under the risk neutral setup, constructs the links between these variables and risk-neutral probabilities. Under the risk neutral setup, the price of the same state contingent asset becomes

\[ S_i = \mathbb{E}^*(S_j) = e^{-r_i} \sum_{j=1}^{m} q_{ij} S_j \quad \text{(3.13)} \]

where \( q_{ij} \)'s are risk-neutral transition probabilities. This equation implies that Arrow-Debreu prices are discounted risk-neutral probabilities. Hence, the relation between risk neutral transition probabilities and Arrow-Debreu prices can be written as

\[ p_{ij} = e^{-r_i} q_{ij} \quad \text{(3.14)} \]
Equation (3.14) implies that we are able to obtain Arrow-Debreu prices as discounted risk-neutral probabilities. We use this representation throughout the paper when we need to compute Arrow-Debreu prices. In a nutshell, we obtain risk neutral probabilities from risk neutral parameters estimated from affine term structure model. Then we convert risk neutral probabilities into Arrow-Debreu prices with the discount factors. Discount factors are obtained as one-period bonds prices within the affine model. The next section elaborates on the details of our approach.

3.1. Implementation of Ross Recovery

The implementation of the Recovery Theorem relies on having Arrow-Debreu prices for all possible state transitions. Thus, the first step in the estimation is to find entries of the matrix \( P \). In a world with three states (Good, Neutral, Bad) this translates to having a matrix with the following entries.

<table>
<thead>
<tr>
<th>State</th>
<th>Good</th>
<th>Neutral</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>( p_{11} )</td>
<td>( p_{12} )</td>
<td>( p_{13} )</td>
</tr>
<tr>
<td>Neutral</td>
<td>( p_{21} )</td>
<td>( p_{22} )</td>
<td>( p_{23} )</td>
</tr>
<tr>
<td>Bad</td>
<td>( p_{31} )</td>
<td>( p_{32} )</td>
<td>( p_{33} )</td>
</tr>
</tbody>
</table>

In this representation \( p_{ij} \) refers to the price of an Arrow-Debreu security moving from state \( i \) to state \( j \). There is a large literature on the estimation of Arrow-Debreu prices from prices of securities traded in the market. However, the attempts in the literature do not produce the representation required by the Recovery Theorem.

The conventional solution to obtain Arrow-Debreu prices is to use -if available- options prices as outlined by Breeden and Litzenberger (1978). In this approach, Arrow-Debreu prices are obtained as second derivative of option prices with respect to exercise price. The tool derived by Breeden and Litzenberger (1978) has been extensively used in the literature. As long as there are sufficient number of options, one can obtain Arrow-Debreu prices using the results of Breeden and Litzenberger (1978).

In our context, Breeden and Litzenberger (1978) approach poses two important limitations. The first limitation is that it is applicable only to options prices while our goal is to apply the Recovery Theorem to bond yields. The second limitation is that with this approach one can estimate Arrow-Debreu prices conditional on the current state of nature because pricing is carried out conditional on current asset price. In other words, Breeden
and Litzenberger (1978) method yield only one row of the Arrow-Debreu price matrix $P$, leaving the questions about other state transitions unanswered.

Despite these limitations, Ross (2015) is able to construct matrix of Arrow-Debreu prices from options prices using Breeden and Litzenberger (1978) approach. He estimates rows of the Arrow-Debreu price matrix using the information that “state price for a security with a payoff in state $j$ at time $t+1$ is the state price for a payment at time $t$ in some intermediate state $k$ multiplied by the transition price of going from state $k$ to state $j$, $p_{kj}$, and then added up over all the possible intermediate states, $k$”.

Central to Ross (2015) analysis is the existence of a large dataset with sufficient maturities of options. When the options prices at different maturities are available, it is possible to obtain multiple rows of Arrow-Debreu matrix $P$ where these rows correspond to states other than current one. Ross (2015) uses a data set obtained from a private bank’s bid price sheet, which is not publicly available. Also, he assumes the number of states to be eleven. These assumptions allow him to construct Arrow-Debreu price matrix from options prices but limit the applicability of his approach to different asset classes such as bond markets.

Another difficulty with the implementation of the Recovery Theorem in different asset classes is the necessity of imposing state-dependent discount factors to obtain physical probabilities that are different from the risk neutral ones. Ross (2015) explains the reasons for this in a corollary. Accordingly, if interest rates are deterministic, the Recovery Theorem suggests that pricing is risk neutral, i.e., the physical and the risk neutral probabilities are the same. Therefore, in order to implement the Recovery Theorem, one needs to impose stochastic interest rates into the pricing functional. When options prices are used, it is possible to circumvent this problem since Breeden and Litzenberger (1978) approach produces Arrow-Debreu prices, implicitly allowing for state-dependence.

The aforementioned issues highlight the important challenges researchers face when applying the Recovery Theorem in different asset classes. We note that the Recovery Theorem is valid for any contingent claims. In this respect, we look for alternative ways of obtaining Arrow-Debreu price matrix. We build our work on by specifying the entries of Arrow-Debreu price matrix, $P$ as product of risk neutral probabilities and corresponding discount factors. In this case, interest rate and risk neutral probability both depend on the states and the matrix entry, $p_{ij}$ becomes

$$p_{ij} = e^{-r_i q_{ij}}$$

(3.15)
Equation (3.15) implies that if we start from the risk neutral pricing, we need to impose an additional stochastic process for modeling discount factors. This structure requires that both risk neutral probabilities and discount factors change with states of nature. As discussed before, this is critical in order to get a nonzero risk premium from the application of the Recovery Theorem. A candidate solution could be to work with two stochastic processes: one for the underlying asset and one for the discount factors (or interest rates) as in Bakshi, Cao, and Chen (1997) for options returns. But this implies that both processes change in line with the states of nature, inducing a correlation between asset prices and discount factors. Construction of Arrow-Debreu prices in this setup might still be problematic if the correlation between asset prices and discount factors does not exist or stay constant. This indicates the need for an approach avoiding such ad-hoc correlations implied by Markov representation of states.

Our solution to this issue is to work with an asset class for which both discount factor and risk neutral transition probabilities are functions of states. We use one period bond prices as discount factors where underlying parameters obtained from bond yields at different maturities. The tractable nature of affine term structure models enables us to obtain both terms from risk-neutral parameters.

4. Econometric Framework

In this section, we explain the Affine Term Structure models as they serve as the building blocks of our estimation strategy. We discuss how they should be used to draw inferences about the expectation and risk premium components embedded in bond yields. We choose discrete time Gaussian model in our specification as outlined by Gürkaynak and Wright (2012). This model captures most of the features of affine models and allows us to establish a link between affine models and the Recovery Theorem.

4.1. Affine Term Structure Model

A common modeling approach in the literature is to assume that there are low dimensional risk factors driving the interest rates. Our model encompasses affine models under the risk neutral measure. We start specifying the risk neutral evolution of the factors. The stochastic
process of these risk factors $X_t$ is assumed to follow a Vector Autoregression:

$$X_{t+1} = \mu^* + \Phi^* X_t + \Sigma \varepsilon_{t+1},$$  \tag{4.1}$$

where the residuals $\varepsilon_{t+1}$ are iid $\mathcal{N}(0, 1)$, the term $\Phi^*$ is the vector of autoregressive coefficients and the term $\Sigma$ denotes the Cholesky decomposition of the variance-covariance matrix i.e., it is a lower triangular matrix\footnote{Dai and Singleton (2000) and Singleton (2009) impose the matrix $\Phi^*$ to be a lower triangular and variance covariance matrix $\Sigma$, to be identity matrix for identification purposes. As it will be clear in the next section when discussing the econometric issues, this is needed if the likelihood function is maximized over all model parameters and when factors are taken hidden. When risk factors are taken as observable factors, we do not need these restrictions. Both autoregressive matrix and variance-covariance matrix can be full matrices.}.

In this representation, the pricing kernel is assumed to be conditionally lognormal. Hence, the form of the kernel is given by the following expression:

$$\exp \left( -r_t - \frac{1}{2} \lambda_t' \Sigma \Sigma' \lambda_t - \lambda_t' \Sigma \varepsilon_{t+1} \right),$$ \tag{4.2}$$

where $\lambda_t$ is the price of risk vector given by

$$\lambda_t = \lambda_0 + \lambda_1' X_t. \tag{4.3}$$

Equations (4.2) and (4.3) characterizes the transformation of risk neutral dynamics into physical ones. The physical dynamics take the form

$$X_{t+1} = \mu + \Phi X_t + \Sigma \varepsilon_{t+1}, \tag{4.4}$$

where the link between physical and risk neutral parameters is given as

$$\Phi = \Phi^* + \Sigma \lambda_1, \tag{4.5}$$

$$\mu = \mu^* + \Sigma \lambda_0. \tag{4.6}$$

The equilibrium condition for the price of an $n$-period bond $P^n_t$ is given as

$$P^n_t = \mathbb{E}_t [M_{t+1} P^{n-1}_{t+1}], \tag{4.7}$$

where $M_{t+1}$ is the stochastic discount factor or the pricing kernel. This relationship holds for all maturities $n$ as long as there is no arbitrage opportunity \cite{Duffie and Kan (1996)}. 

\[\text{14}\]
In this specification, short rate $r_t$ is assumed to be an affine function of the vector of factors $X_t$:

$$ r_t = \delta_0 + \delta_1' X_t. \quad (4.8) $$

Bond prices, $P^n_t$ can be written in exponentially affine form as

$$ P^n_t = \exp \left( A_n + B_n' X_t \right). \quad (4.9) $$

The maturity dependent coefficients $A_n$ are scalars and the term $B_n$ is a $k \times 1$ vector. Coefficients satisfy the familiar Riccati difference equations

$$ A_{n+1} = -\delta_0 + A_n - B_n' \mu^* + \frac{1}{2} B_n' \Sigma \Sigma' B_n, \quad (4.10) $$

$$ B_{n+1} = -\delta_1' + B_n' \Phi^*. \quad (4.11) $$

The difference equations (4.10) and (4.11) can be solved iteratively using starting values $A_1 = -\delta_0$ and $B_1 = -\delta_1'$. Lastly, the model produces bond yields at each maturity that are a function of risk factors, while loadings of these factors change with time to maturity. Accordingly, yield on an $n$-period bond is given as an affine function of risk factors as follows

$$ y^{(n)}_t = -\frac{1}{n} \log(P^n_t) = -\frac{A_n}{n} - \frac{B_n'}{n} X_t \quad (4.12) $$

Affine models have become popular due to their tractability. They represent short rate, market price of risk, and yields at all maturities as affine function of factors. The affine models can be estimated under physical or risk neutral measures using a panel of bond returns. The estimation under risk neutral measure uncovers risk neutral parameters where $\lambda_0 = 0$ and $\lambda_1 = 0$. This result arises from the fact that only risk neutral parameters appear in Equations (4.10) and (4.11). Thus, there is no need to impose the price of risk parameters in affine models to fit them to data since only the risk neutral parameters matter for the estimation.\[3\]

It is worth noting that there is a cost of using a physical model that becomes apparent during the implementation of the model. Specifically, the price of risk parameters increase with the number of underlying factors. The total number of parameters to be estimated

---

\[3\] Cochrane and Piazzesi (2009) build on this observation and follow a similar strategy. They first estimate risk neutral parameters and then incorporate the price of risk parameters separately into their model.
becomes relatively higher in a physical model than in a risk neutral model. For three dimensional risk factors, for example, affine specification requires estimation of 12 more additional parameters since $\lambda_0$ is a 3x1 vector and $\lambda_1$ is a 3x3 matrix of price of risk coefficients. This high number of parameters to be estimated makes the estimation challenging since it requires the maximization of the likelihood function in many dimensions.

Overall, the choice of physical or risk neutral models depends on the purpose of a researcher. If the goal is to estimate parameters for pricing of interest rate instruments, then the risk neutral model gives satisfactory results. However, if the goal is to extract the risk premium and expectations components, the use of physical model is necessary through either by specification of pricing kernel or by the application of the Recovery Theorem as attempted in the current paper.

4.2. Decomposing the Yield Curve

In this section, we discuss two different ways of computing physical expectations to decompose long-term yields into the average expectations of the short rate and the risk premium components. The difference between two approaches stems from the estimation of physical parameters. We briefly review the traditional way of decomposing long term yields from physical parameters, and then elaborate on our Recovery Theorem based approach.

The rationale behind the decomposition of long term yields comes from the expectation hypothesis. In its strong definition, the expectation hypothesis conjectures that long-term bond yields are equal to the average expectation of short rates. Since the expectation hypothesis is rejected in the empirical work, any deviation from expectation hypothesis is attributed to the presence of risk premium. Hence, long-term bond yields are allowed to have two components: average of the short-term rates and risk premium. In mathematical terms the decomposition of long-term bond yields can be expressed as follows

$$y_t^{(n)} = \frac{1}{n} \mathbb{E}_t \left( y_t^{(1)} + y_{t+1}^{(1)} + \ldots + y_{t+n-1}^{(1)} \right) + rpy_t^{(n)}$$

(4.13)

where $y_t^{(1)}$ and $rpy_t^{(n)}$ denote short rate and yield risk premium, respectively.

In the conventional way of decomposition, physical expectation of the short rate is obtained as follows. As documented in the previous section, the physical parameters of the model are estimated by incorporating market price of risk terms, $\lambda_0$ and $\lambda_1$. These terms are embedded into physical evolution of state process (Equation (4.4)) and the estimation is
carried out by maximizing likelihood function\(^4\). In contrast, in the Recovery Theorem based decomposition, the parameters are derived from physical transition probability matrix, which is obtained by the conversion of Arrow-Debreu prices\(^5\).

As a consequence, the decomposition in both approaches requires obtaining physical dynamics. Due to the affine relation assumed between yields and factors, the expectation of one-period ahead short rate is computed as follows:

\[
E_t y_{t+1}^{(1)} = \delta_0 + \delta'_1 E_t X_{t+1} = \delta_0 + \delta'_1 (\mu + \Phi X_t)
\] (4.14)

As it is the physical model, the expectation of short rates for longer periods amounts to the forward iteration of factor dynamics according to Equation (4.4). Therefore, so long as the physical parameters are available, decomposition can be accomplished by a simple recursion of a difference equation.

4.3. The Econometrics of Term Structure Modeling

The traditional way of estimating the model presented in the section \((4.1)\) is to apply maximum likelihood method since likelihood function is available in closed form. However, numerical search is needed to reach the solution of this optimization problem over multiple dimensions. Due to the large number of model parameters and a flat shape of the likelihood function, simultaneous estimation of the all parameters poses econometric challenges in practical implementation of the model.

\(^4\)For models with observable factors, the Ordinary Least Squares may yield physical parameters if there is a separation between risk-neutral and physical parameters as discussed by [Joslin, Singleton, and Zhu (2011)]. We discuss this topic in the next section.

\(^5\)See Appendix for the derivation of model parameters from Markov chain transition probability matrix.
There are a few attempts aiming to alleviate the effects of such econometric obstacles on the estimation of physical parameters. Kim and Orphanides (2012) propose using survey expectations in the estimation of the affine model. The idea is based on incorporating survey expectations of future bond yields as additional observations. The requirement of this approach is that the model is specified in state space representation and filtering methods are employed to estimate all model parameters (including price of risk parameters) simultaneously. Even though this approach results in a greater variation in expectations and produces smaller confidence bands for parameter estimates, the estimation still suffers from the issues in high dimensional optimization problems. Moreover, the survey expectations are not available for all maturities and asset classes. These issues limit the applicability of this approach.

The difficulty of optimization of high dimensional likelihood function has encouraged researchers to look for solutions with which a small number of parameters are to be estimated. Joslin et al. (2011) propose to use observable factors as opposed to unobserved factors in Kim and Orphanides (2012) approach to reach fast and global solutions to the optimization problem. In their work, observable risk factors are formed from any linear combinations of bond yields such as principal components obtained from a panel of yields. Joslin et al. (2011) show that when the factors are chosen as observables, it is possible to represent the model in such a way that there is a separation between the risk neutral and the physical parameters. Therefore, instead of estimating all model parameters simultaneously, one can estimate risk neutral and physical parameters separately and independently from each other.

The estimation strategy proceeds in the following steps. In the first step, an Ordinary Least Squares (OLS) regression is employed to obtain the physical parameters since maximum likelihood can be successfully achieved by OLS for unrestricted vector autoregressions. In the second step, the likelihood function is maximized to find risk neutral parameters using the parameters from the first step as starting values. To reduce computational burden, Joslin et al. (2011) transform affine model into Jordan normalized form so that the number of parameters to be estimated in Equation (4.4) reduces significantly. This makes the task of estimation of risk neutral parameters much easier compared to previous attempts documented in the literature.

The approach suggested by Joslin et al. (2011) provides significant improvement in the estimation of physical parameters. We will follow our discussion by assuming that the factors are selected from the first three principal components as level, slope and curvature factors, following convention.

The convenience provided by Joslin et al. (2011) is due to the rearrangement of constant vector $\mu^*$ and autoregressive feedback matrix $\Phi^*$. In our three factor model this translates into estimating 4 parameters instead of 12 for $\mu^*$ and $\Phi^*$. 

---

6 We will follow our discussion by assuming that the factors are selected from the first three principal components as level, slope and curvature factors, following convention.

7 The convenience provided by Joslin et al. (2011) is due to the rearrangement of constant vector $\mu^*$ and autoregressive feedback matrix $\Phi^*$. In our three factor model this translates into estimating 4 parameters instead of 12 for $\mu^*$ and $\Phi^*$. 

18
estimation of affine term structure models. It has been implemented by Joslin, Priebsch, and Singleton (2014) and Wright (2011) to estimate bond risk premium and expectations components. However Bauer et al. (2014) note that even though the approach brings significant improvement in search for parameters of global optimum, in small samples it produces biased estimates for physical parameters. Therefore, while this approach helps to alleviate the second concern (i.e., flat likelihood function) mentioned by Kim and Orphanides (2012), the first concern (i.e., small sample problem) still remains.

Moreover, the computational advantages of the approach suggested by Joslin et al. (2011) hinge upon the separation property of risk neutral and physical parameters in likelihood construction. The separation property implies that there are no restrictions across physical and risk neutral parameters. If this assumption is violated, all parameters need to be estimated in one step, thus losing the key benefits of Joslin et al. (2011) approach. In this case, even though cross sectional information is still useful for estimating the risk neutral parameters, the time series information is no longer sufficient to uncover physical dynamics. In other words, OLS method cannot be applied to estimate physical parameters when some risks have zero or constant risk premium (i.e. separation property does not hold).

Two studies in the literature providing evidence for the violation of the separation property are Joslin et al. (2014) and Cochrane and Piazzesi (2009). Both studies find that some risks have zero or constant risk premium (i.e. \( \lambda_1 \) in Equation (4.3) includes zeros on the diagonal elements). In our approach, we use the Recovery Theorem to obtain physical parameters from risk neutral counterparts. Whether some risks are priced or not does not affect our conclusions because we obtain physical parameters from risk neutral ones using physical transition probability matrix. Therefore, our approach does not suffer from the violation of separation property.

Before presenting the details of our estimation approach, it is worth mentioning the critique set forth by Bauer et al. (2014). They show that, even if the separation property holds, affine models estimated using Joslin et al. (2014) methodology with small samples yield biased parameters. The bias in the physical parameter estimation has important economic implications for policy analysis in decomposing the expectations and the risk premium components. Specifically, with biased estimates any changes in forward rates are attributed to changes in term premium. This occurs because model implied long-term expectations do

---

8The two papers differ from each other in the parameterization of the market price of risk. In particular, Joslin et al. (2014) propose searching over 219 different risk specification and choose one according to information criteria. In contrast, Cochrane and Piazzesi (2009) stress that excess returns are earned in compensation to changes in level factor. Only two out of twenty risk parameters are allowed to change freely and the other parameters are restricted to zero.
not respond to short-term interest rate changes. In fact, the expectations obtained from small samples of one to two decades in length show no change through time. In light of Federal Reserve’s change of its monetary policy regime during 80’s, any attempts to model interest rates begin only around late 80’s or early 90’s. Bauer et al. (2014) note that such a sample period is not long enough to uncover the true physical parameters. Therefore, policymakers need to develop alternative approaches to alleviate the difficulties associated with the estimation of the risk premium and expectations components. In order to account for this effect, they suggest bias-correction methods for the dynamics of the factors. The approach imposes an increase especially in the persistence of the first factor. Main problem associated with this approach is that it leads to unit root in the factors allowing for interest rates to be negative. Second, the bias correction applies only the first factor for which the bias is highest. But this ignores bias in the other factors even though all factors suffer from the same bias. In our framework, we are able to show that with a persistence less than unity in the first factor, it is possible to reach economically plausible expectations components. In doing so, we do not need any bias correction since the Recovery Theorem allows to obtain high persistence in the factors.

In the empirical section, we estimate our model following Joslin et al. (2011) methodology since it helps us highlight comparative advantages of our approach. Our aim is to show that the Recovery Theorem works even in small samples and that physical parameters can be obtained from risk-neutral parameters without imposing any restrictions on the physical parameters. We first estimate parameters of the risk factors by Ordinary Least Squares. Then using these parameters as starting points, we estimate risk-neutral parameters by maximizing the likelihood function. We then apply the Recovery Theorem and obtain physical parameters once again from risk neutral dynamics. We show the difference between Ordinary Least Squares implied expectations and the Recovery Theorem implied expectations and elaborate on the implications of this difference.

9One remedy is to correct the bias in the estimation following Bauer et al. (2014). However, in our subsequent analysis we will show that bias-corrected expectation component diverges from least squares implied results in large samples for which the bias is small.
4.4. Markov Chain Approximation

In affine term structure models, the factors $X^*_t$ (Equation (4.1)) determine the short term interest rates and hence the discount factors. The evolution of these factors is assumed to follow a multivariate (vector) autoregression. As the first component of Arrow-Debreu prices, we obtain risk neutral transition probabilities from risk factors. We approximate vector autoregressive representation of factors by Markov chains following the convention in economics literature.

The approximation of univariate and/or multivariate autoregressions by Markov chains constitutes a large and well founded literature in economics. It has been used as a remedy to dynamic optimization problems where closed-form solutions are not available. Tauchen (1986) is the pioneering attempt in this direction. We follow Tauchen (1986) approach to approximate a process of the following generic form

$$X_{t+1} = \rho X_t + \varepsilon_{t+1}$$  \hspace{1cm} (4.15)

where $X_{t+1}$ follows a first-order autoregression and $\varepsilon_{t+1}$ is a white noise process with variance $\sigma^2_{\varepsilon}$. The state spaces $\tilde{X}_{t+1}$ for Markov chain takes discrete values, $\tilde{X}^1 < \tilde{X}^2 < \ldots < \tilde{X}^m$.

The state space is designed to cover $k$ standard deviations below and above the mean. Domain of the states is selected as a multiple of the unconditional standard deviation of the process $X_{t+1}$. For univariate processes, the domain is given as $\sigma_X = (\sigma^2_{\varepsilon}/(1 - \rho^2))^{1/2}$. Thus, the lowest and the highest values of the discretized version of this domain (i.e. state space) are $\tilde{X}^1 = -k\sigma_X$ and $\tilde{X}^m = k\sigma_X$, respectively. All remaining states are equally spaced between these two values according to the total number of states. The number of state space $m$ and bandwidth selection parameter $k$ are at user’s control. Of course, the higher the state space, the better the approximation of the process at hand. In sum, application of this method yields Markov states and corresponding transition probabilities between states.

Tauchen (1986) also describes how to approximate multivariate (vector) autoregressions. The difference in the multivariate case appears to be higher dimensional state vector and associated unconditional variance-covariance matrix $\Omega_X$. In a three dimensional vector autoregression, for instance, the process $X_{t+1}$ becomes 3x1 vector and $\Omega_X$ is a 3x3 diagonal matrix. With $m$ states for each factor, the state space has a dimension of $3 \times m^3$. Similar to univariate case, the values of states $\tilde{X}_{t+1}$ are set to minus and plus $k$ times the unconditional standard deviation. For the dimension index $d = \{1, 2, 3\}$, the lowest and the highest values of this state space are $\tilde{X}_d^1 = -k\sigma_{X,d}$ and $\tilde{X}_d^m = k\sigma_{X,d}$, where $\sigma_{X,d}$ is the square root of...
diagonal elements of $\Omega_X$.

Markov chain approximation gives transition probabilities between states. We work with an autoregressive process, and hence the state space generated by the approximation is the set of values that risk factors can attain between $\tilde{X}^1$ and $\tilde{X}^m$. Under affine setting, Markov chain approximation yields states as discrete levels of risk factors and transition probabilities corresponding to these states. Under the risk neutral model, transitions depend on risk neutral parameters. Therefore, transition probabilities obtained by approximation of risk neutral state evolution (Equation (4.1)) produces risk neutral transition probabilities between states.

To sum up, affine model yields explicit solutions to variables we search for the implementation of the Recovery Theorem. Discount factors are obtained as exponentially affine function of state variables. This enables us to rewrite discount factors as functions of states since each state itself is the factor level, $X$, with the transition probability of moving into any other states is available in matrix form. This procedure yields state dependent discount factors, on the one hand and risk neutral probabilities on the other. Multiplication of two terms produces Arrow-Debreu prices. This way, we obtain the unique necessary input, matrix $P$ for the Recovery Theorem.

Our analysis so far lends the answer to our first research question that we pose in Section 1 to be: “How can we apply the Recovery Theorem in term structure models?” We now turn our attention to our second question: “What are the empirical findings if the Recovery Theorem is applied to term structure of bond yields”?

5. Empirical Applications

The three factor term structure model presented in Section 4.1 is estimated using cross section of bond yields. We use the maximum likelihood estimation of Joslin et al. (2011) where factors (three principal components) are assumed to be measured without errors. This method is used by Joslin et al. (2014) and Wright (2011) to decompose risk premium and expectations components with affine term structure models.
5.1. Estimating Nominal Yield Curve

We use annual yields of U.S. Treasury bonds with maturities range from 3-month to 10-years. The dataset includes quarterly observations of zero-coupon government bond yields from 1990:Q1 until 2009:Q1\(^\text{10}\) The sample period is kept relatively short for two reasons. First, we aim to show that the Recovery Theorem based physical parameters yield enough variation in long term interest rates even with small sample periods. Second, as discussed by Kim and Wright (2005) and Wright (2011) the monetary policy regime conducted by Federal Reserve does not show a major change between these dates. However, during 1980’s the main target was to reduce high inflation. Similarly, after the widespread 2008-2009 crisis, FED lowered to interest rates to very low levels to support Large Scale Quantitative Easing. In order to mitigate the effects of these changes on the monetary policy regime and avoid estimation with a potential structural breaks in the data, we follow Wright (2011) and keep the sample between these dates.

We proceed with the estimation using two alternative methods. We first estimate physical dynamics by Ordinary Least Squares as discussed in Joslin et al. (2014). This approach produces physical parameters provided that all risk factors have non-zero price of risk (i.e. separation property holds). In the second approach, we estimate the same model using the Recovery Theorem using the risk-neutral parameters obtained by maximizing the likelihood function.

The estimation outputs of both methods are the state parameters, \((\mu, \Phi)\) and variance-covariance matrix, \(\Sigma\). The parameter estimates and their asymptotic standard errors are given in Table 2 for the OLS model. Inspection of eigenvalues of companion matrix, \(\Phi\) shows persistence of the underlying risk factors. The first two eigenvalues of the companion matrix appear to be 0.92 and 0.82. These values imply high persistence in the factors but they are not close enough to unity to produce sufficient variation in the long-term expectations. This is depicted in Figure 4, which shows expectations component obtained by alternative methods.

The decomposition of long-term yields into the expected short rate and the risk premium is achieved by iterating Equation (4.4) forward. We compute 5 to 10 year expectations for both models following the convention in the literature. The expectations component obtained by the OLS method shows almost constant behavior during the period covering the sample. This issue is discussed by Bauer et al. (2014), Wright (2011) and Joslin et al. (2014). It implies that physical parameters obtained by the OLS does not carry any useful

\(^{10}\)The data is available in the web page of the appendix of Wright (2011).
information for monetary policy purposes since monetary policy target is conducted through managing the long-term expectations.

After the OLS method, we estimate physical parameters by the Recovery Theorem. In this approach, we do not impose any restrictions on the price of risk parameters and hence the separation property is not assumed to hold. We confine ourselves with the econometric convenience (i.e., the reduction in the number of parameters) offered by Joslin et al. (2011) in estimating risk-neutral parameters. Accordingly, the factors are normalized with Jordan decomposition. Then, the model parameters are estimated by numerically maximizing the likelihood function using the OLS parameters as starting values. Estimates of risk neutral parameters produce the following set of parameters in Jordan-normalized form.

$$\mu_{\text{Jordan}}^* = \begin{bmatrix} 0.0027 \\ 0 \\ 0 \end{bmatrix}, \quad \Phi_{\text{Jordan}}^* = \begin{bmatrix} 0.9710 & 0 & 0 \\ 0 & 0.9238 & 0 \\ 0 & 0 & 0.4347 \end{bmatrix}$$

Jordan decomposition provides convenience in both the estimation of parameters and approximation of the state dynamics. It reduces the number of parameters to be estimated by imposing zeros on the second and the third rows of the constant vector and on the off-diagonal elements of the companion matrix, which includes the eigenvalues on the diagonal. The decomposition also simplifies Markov chain approximation since three dimensional state vector is now represented as three univariate autoregressions with possibly correlated errors. The eigenvalues of the risk-neutral companion matrix indicate higher persistence than those obtained by the OLS method. The presence of the persistence in the risk-neutral dynamics requires caution in the approximation since the method suggested by Tauchen (1986) may result in poor approximation in this case[11]. The high persistence has also been encountered in other studies and is a common feature of the risk factors driving the bond yields (Duffee and Stanton (2008)). Under these circumstances, the researcher needs to control for the robustness of the approximation routine. In doing so, we keep the number of states at a level sufficiently high to ensure a good approximation. Then, we attempt to obtain the original parameters from Markov chain matrix as it is the conventional way of checking the approximation error in the literature.

[11] To circumvent this problem, Tauchen and Hussey (1991) suggest use of quadrature based approximation methods. These methods improve approximation results by increasing the fineness of the state space grids. Rouwenhorst (1995) proposes another approximation method that yields even better convergence for persistent processes than the approaches adopted by Tauchen (1986) and Tauchen and Hussey (1991). However, these alternative methods are not available for multivariate autoregressions. Tauchen (1986) approach is the only method that can be used in our context.
The Jordan-normalized form allows vector autoregressions to be written as three univariate autoregressions with errors correlated across equations. It can also be interpreted as a vector autoregression with generalized variance-covariance matrix. Terry and Knotek-II (2011) provide an algorithm to generalize Markov chain approximation suggested by Tauchen (1986) to vector autoregressions with generalized variance-covariance matrix. We use their algorithm as our approximation routine.

We choose the number of states to be 21 and obtain the states and corresponding transition probabilities. We select the bandwidth parameter, $k$ as 3 since probability of having an observation in the tails with a distance more than 3 standard deviation away from the mean is very unlikely for Normally distributed random variables. Thus, the domain of the state space ranges from 3 standard deviations below and above the mean. After obtaining Markov chain matrix, we obtain the parameters from Markov chain as outlined in the Appendix A. The resulting transition probability-induced state dynamics are given as follows

$$
\mu_{Jordan}^* = \begin{bmatrix} 0.0027 \\ 0.0000 \\ 0.0000 \end{bmatrix}, \quad \Phi_{Jordan}^* = \begin{bmatrix} 0.9704 & 0.0002 & 0.0006 \\ 0.0000 & 0.9227 & 0.0001 \\ 0.0000 & 0.0000 & 0.4336 \end{bmatrix}
$$

The robustness check indicates promising accuracy in the approximation. The errors appear to exist after the third decimal points. The literature does not provide a benchmark level of accuracy to test the approximation routines since the approximation accuracy is generally controlled with the number of states. But Terry and Knotek-II (2011) report the accuracy up to two decimal points as sufficient level of accuracy. Our approximation results seem to pass this criterion. As a second check, we also plot approximating Markov chain sequence with the actual data points in the factors in Figures 1. The approximating sequence follows closely the actual realizations for all factors. With these evidences in mind, we assume the level of approximation we have reached to be sufficient for the subsequent analysis. On the other hand, we admit that empirical results obtained from the application of the Recovery Theorem depend on the accuracy of this approximation. Hence, approximation with small number of states may cast doubt on the validity of results obtained with persistent factors.

Having established the discrete state space and risk neutral transition probabilities, we now turn to the construction of Arrow-Debreu prices. Our strategy is to find discount factors

---

12The number of states usually encountered in the literature starts from 7 with no upper limits since increasing the number of states improves approximation performance. But the improvement has limits due to computer precision. We stopped at a point when the increase in the number of states does not produce marginal improvement in the approximation performance.
(i.e. one period bond price) as functions of states. In order to accomplish this task, we use a feature of affine models \([4.12]\) that expresses bond price as a linear function of risk factors. Using risk neutral probabilities and one-period discount factors, we construct Arrow-Debreu prices as product of risk neutral probabilities and discount factors. Thus, we obtain matrix of Arrow-Debreu prices by simple multiplication of two components for each state.

In the next step of the application of the Recovery Theorem, we obtain physical probability transition matrix from Arrow-Debreu prices. We first find Perron root and the eigenvector of the Characteristic Equation \((3.9)\) using the Power method as outlined in Poole (2014). Then, physical transition probabilities are recovered according to Equation \((3.10)\).

Lastly, as we obtain risk-neutral dynamics from risk-neutral transition probabilities, we obtain transition probability-induced autoregressive processes from physical transition probability matrix. This results in the following physical dynamics for the Jordan-normalized factors

\[
\mu_{\text{Jordan}} = \begin{bmatrix}
-0.0008 \\
0.0014 \\
0.0009
\end{bmatrix}, \quad \Phi_{\text{Jordan}} = \begin{bmatrix}
0.9432 & 0.0039 & 0.0152 \\
0.0138 & 0.9217 & -0.0050 \\
0.0031 & -0.0005 & 0.4320
\end{bmatrix}
\]

The effect of the Recovery Theorem shows itself with changes on model parameters. Both constant and autoregressive companion matrix change when we move from risk-neutral to physical world. Off-diagonal elements of the companion matrix seem to be less than 2 percent but different from zero, implying that the investors demand risk premium against correlations among risk factors. The second and the third diagonal elements show minor changes while the first factor drops 3 percent, implying that the risk premium demanded against the level factor more than it is demanded against the slope and the curvature factors. This result is consistent with Cochrane and Piazzesi (2009) that excess returns are earned in compensation to changes in the level risk.

In order to compare the Recovery Theorem based results with those obtained by the OLS method, we transform the Jordan-normalized factors into conventional factors. The Table 3 exhibits the parameters obtained by the Recovery Theorem. Accordingly, The OLS and the Recovery Theorem-implied parameters are similar but the implications of long term expectations differ significantly as we discuss the expectation and the risk premium components.

After obtaining the physical parameters, we compare the implications of the OLS and
the Recovery Theorem based estimations in decomposition of long-term yields. As Figure 4 shows, Least Squares model-implied expectations result in a flat expectations component for future interest rates while the Recovery Theorem-implied expectations show time-varying patterns. The results suggest that the Recovery Theorem-implied expectation component captures more information than suggested by the OLS regressions. Figure 4 also shows the expectations implied by the term structure model used by the Federal Reserve. The Federal Reserve model is based on Kim and Wright (2005) and Kim and Orphanides (2012) such that expectations are updated using survey expectations. The survey expectations of Blue Chip Long Range Survey Forecasts are incorporated into model implied expectations.

5.2. Extensions and Limitations

In this section, we discuss extensions to our approach with the Recovery Theorem. Our first aim is to show that the results obtained by the Recovery Theorem is robust against the choice of bandwidth parameter, $k$. We estimate the same model with a larger state space covering the area of five standard deviation away from the mean. This exercise allows us to uncover the effects of the likelihood of rare events and show the robustness of our approach. Our second aim is to show that our estimation approach can be implemented in different contexts. We change our asset class and determine the yield curve from inflation-indexed yields. The empirical results with inflation-indexed yields confirm our findings for nominal yields. The Recovery Theorem produces variation in long-term expectations when the estimation is carried out with small samples. Our third aim is to lower the number of states in the state space in order to reduce the computational burden. To this aim, we make use of the univariate autoregressive form we reach after Jordan decomposition and find states for individual processes conditioning the other processes at constant values.

Enlarging the State Space

This section discusses the effect of choosing a larger state space in the approximation step on the empirical results. The immediate effect could be to have better approximation to true data generating process since with a larger state space, more extreme observations are explained within the model. On the other hand, this may adversely affect the transformation from risk-neutral to physical world by placing positive probabilities to rare events. Since determining the state space implies ignoring some states of the world, the choice of bandwidth should be relevant to finding the physical parameters.
In search for an answer to this inquiry, we change the bandwidth parameter $k$ from three to five. Then, we follow the same states given in the previous section and obtain physical transition probability matrix. The resulting transition probability matrix produce the following state dynamics for Jordan-normalized factors

$$
\mu_{\text{Jordan}} = \begin{bmatrix}
-0.0050 \\
0.0025 \\
0.0012
\end{bmatrix}, \quad
\Phi_{\text{Jordan}} = \begin{bmatrix}
0.9488 & 0.0020 & 0.0024 \\
0.0083 & 0.9221 & -0.0014 \\
0.0017 & -0.0001 & 0.4334
\end{bmatrix}
$$

The difference in the state dynamics when bandwidth parameter changes from three to five appears to be very small. The diagonal elements are very close to our initial estimates, implying that the persistence of the factors does not change materially. The effect on the expectations and risk premium components is negligible. With a higher state space we obtain very similar expectations component. This is evident in Figure 5, where the expectations component with a larger state space exhibits a pattern which is indistinguishable from its counterpart. These results confirm that our approach is robust to choice of bandwidth parameter.

**Estimating Inflation-Indexed Yield Curve**

In this section we estimate affine model using annual yields of U.S. Treasury Inflation protected securities with maturities range from 5 to 10 years. The dataset includes quarterly observations of zero-coupon government bond yields from 2003:Q1 until 2009:Q1. The data is available in the Federal Reserve as web appendix of Gürkaynak, Sack, and Wright (2010).

As discussed by D’Amico, Kim, and Wei (2010) and Gürkaynak et al. (2010), liquidity component seriously affects TIPS yields. In estimating yield curve dynamics, one needs to take into account liquidity effects in inflation-indexed bonds. However, our aim is to see the applicability of the Recovery Theorem in different contexts. Hence, the sample period is chosen such that the liquidity effects are less prominent in this market. Therefore, we ignore liquidity components and estimate inflation indexed bond yields as real (inflation-adjusted) yields. Since the sample is shorter than for nominal bonds it offers another environment where we can test the Recovery Theorem approach.

The model is in the same spirit as discussed in nominal yield curve. We estimate the physical parameters using Joslin et al. (2011) approach by Ordinary Least Squares and then find estimates based on the Recovery Theorem approach. Due to short sample period and/or
liquidity effects the likelihood optimization stays in local maxima and produces nonstation-
arity for risk-neutral dynamics. This poses a problem because the Recovery Theorem is
applicable only when risk neutral process is stationary. Thus, we estimate the model by im-
posing an assumption of stationarity over our model parameters. This amounts to restricting
the possible parameter space for eigenvalues of $\Phi^*$ to be less than one.

Figure 6 exhibits the expectations components obtained by the Ordinary Least Squares
and the Recovery Theorem. Results indicate findings similar to those we obtain for nom-
inal yield curve. The Least Squares Model-implied expectations indicate slowly changing
behavior whereas the Recovery Theorem-based expectations are time-varying. Tables 3 and
4 demonstrate the differences in parameter estimates for these two models. Main difference
appears to be the persistence of the first factor. In the Recovery Theorem-based model the
first factor is highly persistent and makes the expectation component time-varying.

**Dimension Reduction**

As discussed before, the success of the approximation algorithm lies in the accuracy of
approximation routine and estimated physical probabilities. But approximating three risk
factors simultaneously for vector processes requires very high number of states. For each
factor with $m$ states, we need $m^3$ states for simultaneous approximation of the three-factor
model we use above. For vector processes with more factors, building transition matrix
appears computationally challenging. In order to circumvent this challenge, we approximate
each factor separately using the Jordan normalized form obtained in model estimation. In the
approximation step, we use the same estimated parameters obtained for risk neutral process.
But, we approximate the Jordan normalized factor, $X_{1t}$ by conditioning on other factors (i.e.
keeping other factors at population or steady-state means). Similarly, we approximate the
second and the third factors conditioning on the other two factors.

The literature on the Markov chain approximation shows that higher number of states
lead to better approximation for processes. We start increasing the number of states to
get high precision in the approximation. However, the rise in the number of states create
additional complications in the analysis. We observe that there is a trade-off between in-
creasing the number of states and reducibility of the Arrow-Debreu matrix. As we increase
the number of states, the matrix becomes sparse and reducible with many zeros. However,
the Recovery theorem is developed on the assumption that Arrow-Debreu price matrix is
irreducible since Perron-Frobenius theorem is valid only for irreducible matrices. In other
words, increasing the number of states reckless undermine the applicability of our approach
by producing a transition matrix in which jumping from current state to some states are not possible. Economically, this is equivalent to having a complete market which allows investors to insure against all risks associated with the states of the world. Thus, reducibility creates incomplete market structure and the Recovery Theorem becomes inapplicable. In sum, we have an upper limit on the number of states such that higher numbers lead to irreducible matrices. For this reason, we keep the number of states as 31. This way, we guarantee that Perron root exists for the Arrow-Debreu matrix.

The results with this second approach indicate findings similar to those we obtain with our initial application of the Recovery Theorem. Derived expectations component is depicted in Figure 5. This approach allows us to work with more factors within our framework.

Limitations

Our first limitation is related to the specification of affine term structure models. Among many alternatives, we choose discrete and Gaussian model as our building block. However, there is overwhelming evidence in the literature that bond yields are conditionally heteroskedastic. Hence extreme observations in our sample should capture more information about risk-neutral dynamics than Gaussian model implies. As our aim is to see the effects of the Recovery Theorem in the applications, we confine ourselves with the simplest version of affine models. Moreover, Markov chain approximation routine developed by Terry and Knotek-II (2011) is available for vector autoregressions with multivariate normal error structure. Introducing stochastic volatility effects to allow for heteroskedasticity changes normality of the error distribution. Thus, our approach is not applicable for models with conditionally heteroskedastic bond returns.

The second limitation of our model is the stationarity requirement in the risk-neutral data generating process. This is necessary to apply the Markov chain approximation to risk-neutral dynamics. Even though the estimated factor dynamics from nominal yield curve appear to be stationary, those obtained from inflation-indexed yields are not stationary. We can impose stationarity during the likelihood optimization by constraining the set the parameters can reach. But there is no economic reason to impose such constraints on parameters. Nevertheless, this is a drawback of the Recovery Theorem rather than solely of our approach since the Recovery Theorem requires stationarity in risk-neutral dynamics.

Last but not least, our approach does not encompass all models with statistical and/or observable risk factors. If there are macroeconomic or statistical factors affecting bond
yields, we can incorporate these factors into our model as additional factors as carried out by Cochrane and Piazzesi (2009) and Wright (2011). On the other hand, whether the factors span the cross section of bond yields or not makes a difference in the applications. This issue is first raised by Joslin et al. (2014) in disentangling the factors that span the yield curve from those that do not. Accordingly, only the factors spanning the yield curve should be included in the estimation of risk-neutral parameters. This implies that the number of factors in the risk-neutral state dynamics could be different from the number of factors in the physical state dynamics. Our approach goes in line with Joslin et al. (2014) by using only spanning factors in the analysis. But like others it suffers from the same critique if there other factors affecting physical dynamics. We conclude that our approach does not account for such differences in the state dynamics.

6. Business Cycle Implications

In this section, we explore information content of our estimation outputs. Our estimation outputs are the expectation and risk premium components for long-term bond yields. We aim to uncover the implications of these components in policy making. One limitation is that we are unable to test whether expectation and/or risk premium components are true estimates since there is no benchmark serving to this aim. Rather, we would like to see the relative contribution of our estimates compared to other alternatives. To accomplish this task, we investigate the relation of these components to business cycle variables. In particular, we empirically test whether risk premium has a predictive content beyond the expectations components over business cycles.

Making inferences about economic activity requires longer samples than those that we employed in previous sections. We first estimate the nominal term structure model using data covering 1971:Q4-2009:Q1. We obtain the OLS and the Recovery Theorem-implied expectation and risk premium components. Using a longer sample mitigates the bias in the estimation of factor dynamics. Therefore, we expect lower bias associated with the OLS model. This allows the OLS model to serve as a benchmark for the expectations components despite the evidence against the separation property and possible structural breaks in the sample period. The results are depicted in Figure 7 where the Recovery Theorem-based

\[\text{Cochrane and Piazzesi (2009) introduce return forecasting factor and include it as the fourth factor within affine class. Similarly, Wright (2011) determines the number of factors as five by adding inflation and growth factors. In both papers, the factors follow the same autoregressive form. Hence, expectations and risk premium can be estimated using our model.}\]
and Least Squares estimations show close associations. We also plot Bias-corrected model ([Bauer et al., 2014]) estimates to show the relative performance of all alternative models. As is evident from the results, the bias-corrected model based expectations diverge from the others in certain periods. Even though the sample may not be long enough to remove the bias completely from the Least Squares regressions, it is clear that the Recovery Theorem model shows more association with the Least Squares model.

In order to test the importance of risk premium over business cycles, we decompose term spread into the expectations and risk premium components. The literature on information content of term structure includes many attempts showing the significance of term spread in explaining economic activity, where term spread is defined as the difference between long-term and short-term yields. But the marginal contribution of the expectation and risk premium components provides additional clues about the relation between yield curve and economic activity. This is because term spread may signal recessions whenever yield curve is inverted but this may not materialize in reality ([Rosenberg and Maurer, 2008]).

Previous studies such as [Rosenberg and Maurer, 2008], [Ang, Piazzesi, and Wei, 2006], [Wright, 2006] and [Rudebusch et al., 2007] find that expected component is the main driver of the power of term spread in forecasting recessions. Regression exercises in these papers show that risk premium component is insignificant when regressions are performed using GDP growth or recession indicators. Therefore, risk premium should be removed from term spread when used in predicting economic activity. [Hamilton and Kim, 2002] find a significant risk premium coefficient but add that they are not able to find a theoretical justification for this result. [Cochrane, 2007] states that the risk premium should not be significant in forecast regressions but should exhibit counter cyclical pattern as in [Cochrane and Piazzesi, 2005]. As a result, previous results are in favor of using expectation component-only model in forecast regressions but risk premium should rise during the recessions. For this reason, FED officials use both expectation-only and spread models in assessing bond market developments ([Rosenberg and Maurer, 2008]).

As we obtain risk premium from the difference of long term yields and expected yields, high values of risk premium imply that the market expects lower interest rate in the future. This could be an indicator of loose monetary policy and hence future recessions. To investigate this relation further, we plot risk premium estimates obtained in the previous section. Figure 8 presents the National Bureau of Economic Research (NBER) recession indicators.

---

14 The reason of this divergence seems to be excessive persistence generated by bias-correction for the first factor in [Bauer et al., 2014]. Our empirical exercises show that in most cases bias-correction leads to unit-root for the first factor as acknowledged in their work as a limitation of bias-correction.
together with our risk premium estimates. It is evident that risk premium takes high values at the onset of recessions. Following recessions, there is a downward trend in risk premium indicating that expectations get close to current yields in these periods. The pattern of risk confirms the theoretical and empirical results documented by Bauer et al. (2014).

We undertake two regression exercises: one with recessions and one with GDP growth. We first employ Probit regressions to predict NBER recessions. The model is specified as follows

\[
P\left(NBER_{t,t+4} = 1\right) = \Phi\left(\alpha + \beta \text{Spread}_t\right)
\]

where \(NBER_{t,t+4}\) is the dummy that takes on value 1 if there is an NBER defined recession at some point during the quarters from \(t + 1\) to \(t + 4\) and \(\Phi(.)\) denotes the standard normal cumulative distribution function.

The model in Equation (6.1) follows closely the works of Estrella and Hardouvelis (1991) but differs from this study in that our aim is to see the effect of each component on recession predictions as Wright (2006) and Rudebusch et al. (2007). The significance of individual terms gives us clues about the validity of our decomposition.

We estimate the Probit model with maximum likelihood. Due to overlapping periods in errors, we report robust standard errors\(^{15}\). The results are presented in Table 6.

The estimation results show that both the expectation and risk premium components obtained from the Least Squares model are significant in predicting recessions. But the tests for the equality of coefficients are in favor of restricted model suggesting spread-only model. This implies that both components carry information about future activity, and that the spread model should capture the information about recessions. However, when the decomposition is done according to the Recovery Theorem-based model, the expectation component appears to be significant while risk premium component is not. This suggests that risk premium should be removed from the model specification when forecasting recessions. These results imply that our decomposition exercise produces similar results to previous studies.

Due to persistent behavior in risk premium and expectations components, we also run the regressions with four period changes in the variables. This approach is suggested by

\(^{15}\)Estrella and Rodrigues (1998) show that Hansen (1982) yields the best covariance matrix for Probit regressions when errors are autocorrelated. We follow their definition of robust standard errors in our analysis.
Rudebusch et al. (2007). Accordingly, not the level but the changes in risk premium become of primary relevance from the monetary policy perspective. Our results confirm their findings that changes in risk premium obtained from the Recovery Theorem model is significant while the changes in the expectation component are not. In all models, the Recovery Theorem based model has higher $R^2$ suggesting an increase in predictive power. We perform the same exercise using 6-quarter ahead recession indicator and obtain results that are qualitatively the same. The results are presented in Table 7.

As a second exercise, we run forecasting regressions using GDP growth. Since recessions are rare events, it is better to see the relevance of our derived measures in economic activity. The model follows Rudebusch et al. (2007)

\[
y_{t,t+4} = \alpha + \beta_1 y_{t-4,t} + \beta_2 \text{Spread}_t + \epsilon_t
\]

(6.2)

where $y_{t,t+4}$ denote the 4-period ahead growth rate in GDP and $y_{t-4,t}$ is the 4-period realized growth in GDP. The model helps uncover the effect of bond risk premium in GDP forecast regressions. The empirical results with the growth variable lead to the same conclusion we reached in Probit regressions. Table 8 presents the forecast regressions with GDP growth. The results indicate that for the expectations component obtained by the Recovery Theorem, the level of the component is significant. But for the risk premium component, only 4-quarter changes are significant.

7. Conclusion

This paper attempts to decompose bond yields into the expectation and risk premium components using Ross (2015) Recovery Theorem. Our strategy is to utilize affine term structure models and obtain necessary inputs for Ross’s Recovery Theorem. We are able to disentangle these two components without specifying the market price of risk.

The recovery model is applicable in short samples where alternative models require either the use of additional data such as surveys or bias corrections. Our model allows us to obtain precise estimates of physical parameters and get expectations and risk premium components depending solely on the information in bond yields.

Derived expectations component shows substantial variation over time, while the difference between long-term yields and expectations captures information about variation of risk through time. The Recovery Theorem-based risk premium and expectations components
have similar forecasting power over recessions and GDP growth to its alternatives. Risk pre-
mium follows a counter cyclical pattern rising in the early phase of recessions. Also, changes
in risk premium are significant in forecast regressions for economic activity.

Our model has appealing features encapsulating the information content of the term
structure of bond yields. It can be used for policy analysis and investment decisions. We
show two examples where estimated components can be used in policy making. We find
that the expectation component has strong predictive power for future recessionary periods.
Specifically, it predicts recessions 4 to 6 quarters ahead of the official announcement of
recessions.
Appendix A. Derivation of State Parameters

This appendix provides the formulas needed to obtain constant $\mu$ and feedback matrix $\Phi$ parameters in state evolution (Equation 4.4). The solution is provided for a given physical Markov Chain transition probability matrix $F$.

A.1. Stationary Probability Distribution

The stationary probability distribution $\lambda$ of the Markov chain transition matrix $F$ can be found by solving the following characteristic equation. The stationary distribution corresponds to the eigenvector of the first eigenvalue which is less than or equal to maximum row sum according to Perron-Frobenius theorem. Since this is a probability matrix, maximum row sum is 1.

$$\lambda = \lambda F. \tag{A.1}$$

A.2. Conditional Mean

Conditional expectation of the discretized process $\tilde{X}_{t+1}$ can be found from the transition probability matrix $F$. If the state space takes values $\tilde{X} = \{x_1, x_2, ..., x_j, ..., x_m\}$ with probability $f_{ij}$ then conditional expectation is given as

$$\mathbb{E}(\tilde{X}_{t+1} | \tilde{X}_t = x_i) = \bar{x}_i = \sum_{j=1}^{m} f_{ij} x_j \tag{A.2}$$

A.3. Unconditional Mean

Unconditional mean of the process $\tilde{X}$ can be found from the stationary probabilities $\lambda$ as follows.

$$\mathbb{E}(\tilde{X}_{t+1}) = \bar{x} = \sum_{j=1}^{m} \lambda_j x_j \tag{A.3}$$
A.4. Unconditional Variance

Unconditional variance of the process $\tilde{X}$ can be found from the stationary probabilities $\lambda$ as follows.

$$\mathbb{E}\left(\left[\tilde{X}_{t+1} - \mathbb{E}(\tilde{X}_{t+1})\right]^2\right) = \sum_{j=1}^{m} \lambda_j (x_j - \bar{x})^2$$  \hspace{1cm} (A.4)

A.5. Unconditional Covariances

The covariances between individual processes $\tilde{X}_{d,t+1}$, where $d$ denotes dimension index, can be obtained from the stationary probabilities $\lambda$ and unconditional mean $\bar{x}^d$ of the same process as follows.

$$\mathbb{E}\left(\left[\tilde{X}_{1,t+1} - \mathbb{E}(\tilde{X}_{1,t+1})\right]\left[\tilde{X}_{2,t+1} - \mathbb{E}(\tilde{X}_{2,t+1})\right]\right) = \sum_{j=1}^{m} \lambda_j (x_{1,j} - \bar{x}^1)(x_{2,j} - \bar{x}^2)$$  \hspace{1cm} (A.5)

These moments enable us to find the all combinations of multiplications of individual processes $\tilde{X}_{d,t+1}$. Using the information that Markov chain probabilities do not change in time (time-homogeneity), we form the following expectations terms.

$$\mathbb{E}(\tilde{X}_{1,t+1}^2) = \mathbb{E}(\tilde{X}_{2,t+1}^2) = \sum_{j=1}^{m} \lambda_j x_{1,j}^2$$ \hspace{1cm} (A.6)

$$\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{2,t+1}) = \mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{2,t+1}) = \sum_{j=1}^{m} \lambda_j x_{1,j} x_{2,j}$$ \hspace{1cm} (A.7)

Similarly the other covariance terms are obtained from \([A.2]\) conditional and \([A.3]\) unconditional.
ditional means.

\[
\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{1,t}) = \mathbb{E}\left(\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{1,t}|\tilde{X}_{1,t} = x_{1,j})\right)
\]

\[
= \sum_{j=1}^{m} \lambda_{j}x_{1,j}\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{1,t} | \tilde{X}_{1,t} = x_{1,j})
\]

\[
= \sum_{j=1}^{m} \lambda_{j}x_{1,j}\sum_{k=1}^{m} f_{jk}x_{1,j}
\]

\[
= \sum_{j=1}^{m} \lambda_{j}x_{1,j}\bar{x}_{1,j}
\]

(A.8)

\[
\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{2,t}) = \mathbb{E}\left(\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{2,t}|\tilde{X}_{2,t} = x_{2,j})\right)
\]

\[
= \sum_{j=1}^{m} \lambda_{j}x_{2,j}\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{2,t} | \tilde{X}_{2,t} = x_{2,j})
\]

\[
= \sum_{j=1}^{m} \lambda_{j}x_{2,j}\sum_{k=1}^{m} f_{jk}x_{1,j}
\]

\[
= \sum_{j=1}^{m} \lambda_{j}x_{2,j}\bar{x}_{1,j}
\]

(A.9)

Lastly, we are able construct the following matrices.

\[
\mathcal{E} = \begin{bmatrix}
1 & \mathbb{E}(\tilde{X}_{1,t}) & \mathbb{E}(\tilde{X}_{2,t}) & \mathbb{E}(\tilde{X}_{3,t}) \\
\mathbb{E}(\tilde{X}_{1,t}) & \mathbb{E}(\tilde{X}_{1,t}^2) & \mathbb{E}(\tilde{X}_{1,t}\tilde{X}_{2,t}) & \mathbb{E}(\tilde{X}_{1,t}\tilde{X}_{3,t}) \\
\mathbb{E}(\tilde{X}_{2,t}) & \mathbb{E}(\tilde{X}_{1,t}\tilde{X}_{2,t}) & \mathbb{E}(\tilde{X}_{2,t}^2) & \mathbb{E}(\tilde{X}_{2,t}\tilde{X}_{3,t}) \\
\mathbb{E}(\tilde{X}_{3,t}) & \mathbb{E}(\tilde{X}_{1,t}\tilde{X}_{3,t}) & \mathbb{E}(\tilde{X}_{2,t}\tilde{X}_{3,t}) & \mathbb{E}(\tilde{X}_{3,t}^2)
\end{bmatrix}
\]

\[
\mathcal{X}_1 = \begin{bmatrix}
\mathbb{E}(\tilde{X}_{1,t}) \\
\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{1,t}) \\
\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{2,t}) \\
\mathbb{E}(\tilde{X}_{1,t+1}\tilde{X}_{3,t})
\end{bmatrix}, \quad \mathcal{X}_2 = \begin{bmatrix}
\mathbb{E}(\tilde{X}_{2,t}) \\
\mathbb{E}(\tilde{X}_{2,t+1}\tilde{X}_{2,t}) \\
\mathbb{E}(\tilde{X}_{2,t+1}\tilde{X}_{2,t}) \\
\mathbb{E}(\tilde{X}_{2,t+1}\tilde{X}_{2,t})
\end{bmatrix}, \quad \mathcal{X}_3 = \begin{bmatrix}
\mathbb{E}(\tilde{X}_{3,t}) \\
\mathbb{E}(\tilde{X}_{3,t+1}\tilde{X}_{3,t}) \\
\mathbb{E}(\tilde{X}_{3,t+1}\tilde{X}_{3,t}) \\
\mathbb{E}(\tilde{X}_{3,t+1}\tilde{X}_{3,t})
\end{bmatrix}
\]
The parameter values are given as the following algebraic solutions.

\[
\begin{bmatrix}
\mu_1 \\
\Phi_{11} \\
\Phi_{12} \\
\Phi_{13}
\end{bmatrix} = \mathcal{E}^{-1} \mathcal{X}_1, \\
\begin{bmatrix}
\mu_2 \\
\Phi_{21} \\
\Phi_{22} \\
\Phi_{23}
\end{bmatrix} = \mathcal{E}^{-1} \mathcal{X}_2, \\
\begin{bmatrix}
\mu_3 \\
\Phi_{31} \\
\Phi_{32} \\
\Phi_{33}
\end{bmatrix} = \mathcal{E}^{-1} \mathcal{X}_3.
\]
References


40


URL http://dx.doi.org/10.1111/jofi.12092


Fig. 1. Approximation of factors: Each factor is approximated with a discrete Markov chain. Corresponding states of the transition matrix for a given level of a factor are shown in red circles. The sample covers the period Q1:1990-Q1:2009.
Fig. 2. Approximation of factors: Each factor is approximated with a discrete Markov chain. Corresponding states of the transition matrix for a given level of a factor are shown in red circles. The sample covers the period Q1:2003-Q1:2009.
Fig. 3. Approximation of factors: Each factor is approximated with a discrete Markov chain. Corresponding states of the transition matrix for a given level of a factor are shown in red circles. The sample covers the period Q4:1971-Q1:2009.
Fig. 4. This figure plots expectations of five-to-ten-year ahead short rates obtained from three different models: Ordinary Least Squares, the Recovery Theorem, and the model maintained by the Federal Reserve Board based on (Kim and Wright 2005; and Kim and Orphanides 2012). The sample period is Q1:1990-Q1:2009.
Fig. 5. This figure plots expectations of five-to-ten-year ahead short rates obtained from three different versions of the Recovery Theorem. The baseline model has a state space covering three standard deviation below and above the mean. Enlarged state space model covers five standard deviation around the mean. Reduced dimension model forms the state space by using the univariate representation of Jordan-normalized form and conditions other states at population means. Sample period is Q1:1990-Q1:2009.
Fig. 6. This figure plots expectations of five-to-ten-year ahead real rates obtained from Ordinary Least Squares estimation and from the Recovery Theorem. The sample period is Q1:2003-Q1:2009.
Fig. 7. This figure plots expectations of five-to-ten-year ahead short rates obtained from three different models: Ordinary Least Squares, the Recovery Theorem and the Bias-corrected Least Squares. The sample period is Q4:1971-Q1:2009.
Fig. 8. This figure plots five-to-ten-year ahead risk premium estimates obtained from the Recovery Theorem. Shaded regions denote NBER defined recession periods. The sample period is Q4:1971-Q1:2009.
Table 2: Estimated parameters for the nominal yield curve using the sample covering Q1:1990-Q1:2009. Asymptotic standard errors are given in parentheses.
Estimated Parameters: Recovery model

<table>
<thead>
<tr>
<th></th>
<th>0.0002</th>
<th>0.0000</th>
<th>0.0001</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0004)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>1.0020</td>
<td>0.1590</td>
<td>-0.8905</td>
</tr>
<tr>
<td></td>
<td>(0.0045)</td>
<td>(0.0160)</td>
<td>(0.0842)</td>
</tr>
<tr>
<td></td>
<td>-0.0215</td>
<td>0.9930</td>
<td>0.8733</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0051)</td>
<td>(0.0267)</td>
</tr>
<tr>
<td></td>
<td>0.0091</td>
<td>-0.0756</td>
<td>0.3020</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td>(0.0018)</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>$\Sigma x 10^3$</td>
<td>0.4070</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5756)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0627</td>
<td>0.0410</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1436)</td>
<td>(0.0580)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0333</td>
<td>-0.0070</td>
<td>0.0053</td>
</tr>
<tr>
<td></td>
<td>(0.0573)</td>
<td>(0.0164)</td>
<td>(0.0075)</td>
</tr>
</tbody>
</table>

Table 3: Estimated parameters for the nominal yield curve using the sample covering Q1:1990-Q1:2009. The state space is selected as three standard deviation above and below the mean. Asymptotic standard errors are given in parentheses.
### Estimated Parameters: Least Squares model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Standard Error</th>
<th>Z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.0552</td>
<td>-0.0088</td>
<td>-0.0003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0242)</td>
<td>(0.0034)</td>
<td>(0.0005)</td>
<td></td>
</tr>
<tr>
<td>$\Phi$</td>
<td>0.2558</td>
<td>0.6009</td>
<td>5.1324</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2070)</td>
<td>(0.7883)</td>
<td>(9.9074)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0452</td>
<td>0.8173</td>
<td>1.8671</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0290)</td>
<td>(0.1105)</td>
<td>(1.3890)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0066</td>
<td>-0.0096</td>
<td>0.7327</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0046)</td>
<td>(0.0175)</td>
<td>(0.2196)</td>
<td></td>
</tr>
<tr>
<td>$\Sigma x10^3$</td>
<td>0.2173</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3074)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0109</td>
<td>0.0043</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0323)</td>
<td>(0.0060)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0006</td>
<td>-0.0001</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0049)</td>
<td>(0.0007)</td>
<td>(0.0002)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Estimated parameters for the inflation-indexed yield curve using the sample covering Q1:2003-Q1:2009. Asymptotic standard errors are given in parentheses.
<table>
<thead>
<tr>
<th>Estimated Parameters: Recovery model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\Phi$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\Sigma \times 10^3$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Table 5: Estimated parameters for the inflation-indexed yield curve using the sample covering Q1:2003-Q1:2009. Asymptotic standard errors are given in parentheses.
<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$spr_t$</td>
<td>-0.569***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.496)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$rP_{t}^{ols}$</td>
<td>-0.606***</td>
<td>-0.562</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.195)</td>
<td>(-1.436)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e_x^{ols}$</td>
<td>-0.562***</td>
<td>-0.568***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.347)</td>
<td>(-4.500)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$rP_{t}^{rec}$</td>
<td></td>
<td></td>
<td>1.391***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.590)</td>
<td></td>
</tr>
<tr>
<td>$e_x^{rec}$</td>
<td></td>
<td></td>
<td>-0.144</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-1.367)</td>
<td></td>
</tr>
<tr>
<td>$\Delta rP_{t}^{ols}$</td>
<td>-0.374**</td>
<td>-0.319</td>
<td>-0.385</td>
<td>-1.017***</td>
</tr>
<tr>
<td></td>
<td>(-1.997)</td>
<td>(-1.176)</td>
<td>(-0.639)</td>
<td>(-7.901)</td>
</tr>
<tr>
<td>$\Delta e_x^{ols}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta rP_{t}^{rec}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>(-1.997)</td>
<td>(-1.176)</td>
<td>(-0.639)</td>
<td>(-7.901)</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.251</td>
<td>0.252</td>
<td>0.251</td>
<td>0.0270</td>
</tr>
<tr>
<td>Wald-test ($e_x = yp_t$)</td>
<td>0.068</td>
<td>2.96e-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald-test ($\Delta e_x = \Delta yp_t$)</td>
<td></td>
<td></td>
<td>2.670</td>
<td>7.581</td>
</tr>
<tr>
<td>Prob $&gt; F$</td>
<td>0.795</td>
<td>0.986</td>
<td>0.102</td>
<td>5.96e-03</td>
</tr>
</tbody>
</table>

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: This table reports Probit regression results for the NBER defined recessions using term spread, $spr_t$ obtained as the difference between 10-year and 3-month interest rates and its decomposition. $rP_{t}^{ols}$ and $e_x^{ols}$ denote risk premium and expectation components from the Ordinary Least Squares regression while $rP_{t}^{rec}$ and $e_x^{rec}$ denote the same components obtained from the Recovery Theorem model. $\Delta$ denotes 4-quarter changes in the variables. $rec_{t+4}$ takes the value 1 if there is a recession within 4 quarters ahead.
### Table 7

This table reports Probit regression results for the NBER defined recessions using term spread, $spr_t$, obtained as the difference between 10-year and 3-month interest rates and its decomposition. $r_{p_{t}}^{ols}$ and $e_{x_{t}}^{ols}$ denote risk premium and expectation components from the Ordinary Least Squares regression while $r_{p_{t}}^{rec}$ and $e_{x_{t}}^{rec}$ denote the same components obtained from the Recovery Theorem model. $\Delta$ denotes 4-quarter changes in the variables. $rec_{t+6}$ takes the value 1 if there is a recession within 6 quarters ahead.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$spr_t$</td>
<td>-0.581***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.957)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{p_{t}}^{ols}$</td>
<td></td>
<td>-0.685***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-4.126)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e_{x_{t}}^{ols}$</td>
<td></td>
<td></td>
<td>-0.566***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-4.784)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{p_{t}}^{rec}$</td>
<td></td>
<td></td>
<td>-0.906**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-2.430)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e_{x_{t}}^{rec}$</td>
<td></td>
<td></td>
<td>-0.601***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-5.194)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta r_{p_{t}}^{ols}$</td>
<td></td>
<td></td>
<td></td>
<td>0.266</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.536)</td>
<td></td>
</tr>
<tr>
<td>$\Delta e_{x_{t}}^{ols}$</td>
<td></td>
<td></td>
<td></td>
<td>-0.465***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-3.547)</td>
<td></td>
</tr>
<tr>
<td>$\Delta r_{p_{t}}^{rec}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.590**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.544)</td>
</tr>
<tr>
<td>$\Delta e_{x_{t}}^{rec}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.288**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-2.560)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.377*</td>
<td>-0.232</td>
<td>0.115</td>
<td>-1.135***</td>
<td>-1.157***</td>
</tr>
<tr>
<td></td>
<td>(-1.938)</td>
<td>(-0.939)</td>
<td>(0.205)</td>
<td>(-8.093)</td>
<td>(-8.150)</td>
</tr>
<tr>
<td>Observations</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>146</td>
<td>146</td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.251</td>
<td>0.257</td>
<td>0.258</td>
<td>0.109</td>
<td>0.125</td>
</tr>
<tr>
<td>Wald-test ($e_{x_{t}} = yp_{t}$)</td>
<td>0.682</td>
<td>0.785</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald-test ($\Delta e_{x_{t}} = \Delta yp_{t}$)</td>
<td></td>
<td></td>
<td>10.59</td>
<td>8.190</td>
<td></td>
</tr>
<tr>
<td>Prob &gt; F</td>
<td>0.409</td>
<td>0.376</td>
<td>1.14e-03</td>
<td>4.21e-03</td>
<td></td>
</tr>
</tbody>
</table>

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Table 8: This table reports forecasting regression results for the GDP using term spread, \( spr_t \) obtained as the difference between 10-year and 3-month interest rates and its decomposition. \( r_{p_t}^{ols} \) and \( e_{x_t}^{ols} \) denote risk premium and expectation components from the Ordinary Least Squares regression while \( r_{p_t}^{rec} \) and \( e_{x_t}^{rec} \) are the same components obtained from the Recovery Theorem model. \( \Delta \) denotes 4-quarter changes in the variables.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( y_{t+4} - y_t )</td>
<td>( y_{t+4} - y_t )</td>
<td>( y_{t+4} - y_t )</td>
<td>( y_{t+4} - y_t )</td>
<td>( y_{t+4} - y_t )</td>
</tr>
<tr>
<td>( y_{t-4} - y_t )</td>
<td>0.218</td>
<td>0.210*</td>
<td>0.218*</td>
<td>0.303**</td>
<td>0.318**</td>
</tr>
<tr>
<td></td>
<td>(1.648)</td>
<td>(1.711)</td>
<td>(1.675)</td>
<td>(2.487)</td>
<td>(2.491)</td>
</tr>
<tr>
<td>( spr_t )</td>
<td>0.648***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.990)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_{p_t}^{ols} )</td>
<td></td>
<td>0.815***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.876)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( e_{x_t}^{ols} )</td>
<td></td>
<td>0.595***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.124)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_{p_t}^{rec} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.820</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.104)</td>
</tr>
<tr>
<td>( e_{x_t}^{rec} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.646***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(3.954)</td>
</tr>
<tr>
<td>( \Delta r_{p_t}^{ols} )</td>
<td></td>
<td></td>
<td></td>
<td>-0.628*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-1.832)</td>
<td></td>
</tr>
<tr>
<td>( \Delta e_{x_t}^{ols} )</td>
<td></td>
<td></td>
<td></td>
<td>0.619**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.018)</td>
<td></td>
</tr>
<tr>
<td>( \Delta r_{p_t}^{rec} )</td>
<td></td>
<td></td>
<td></td>
<td>-3.429***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-2.802)</td>
<td></td>
</tr>
<tr>
<td>( \Delta e_{x_t}^{rec} )</td>
<td></td>
<td></td>
<td></td>
<td>0.359*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.781)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.133*</td>
<td>0.888</td>
<td>0.856</td>
<td>1.824***</td>
<td>1.800***</td>
</tr>
<tr>
<td></td>
<td>(1.890)</td>
<td>(1.315)</td>
<td>(0.647)</td>
<td>(3.408)</td>
<td>(3.329)</td>
</tr>
<tr>
<td>Observations</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>146</td>
<td>146</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.186</td>
<td>0.196</td>
<td>0.187</td>
<td>0.131</td>
<td>0.187</td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.175</td>
<td>0.179</td>
<td>0.170</td>
<td>0.113</td>
<td>0.169</td>
</tr>
<tr>
<td>F-test (( e_{x_t}=y_{p_t} ))</td>
<td>0.416</td>
<td>0.053</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-test (( \Delta e_{x_t}=\Delta y_{p_t} ))</td>
<td></td>
<td></td>
<td>5.568</td>
<td>8.64</td>
<td></td>
</tr>
<tr>
<td>Prob &gt; F</td>
<td>0.520</td>
<td>0.818</td>
<td>0.020</td>
<td>0.004</td>
<td></td>
</tr>
</tbody>
</table>

* **p<0.01, * p<0.05

t-statistics adjusted by Newey-West standard errors in parentheses.