The Term Structure of Interest Rates and Monetary Policy: An Empirical Assessment

Inaugural-Dissertation zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaft
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

Freie Universität Berlin

vorgelegt von

Till Strohsal

aus Berlin

im August 2013

Erstgutachter:	Prof. Dr. Dieter Nautz
Zweitgutachter:	Prof. Dr. Helmut Lütkepohl
Tag der Disputation:	9. Dezember 2013

Eigenanteil der Leistung

Diese Dissertation besteht aus vier Arbeitspapieren, von denen drei in Zusammenarbeit mit jeweils einem Koautor entstanden sind. Der Eigenanteil an Konzeption, Durchführung und Berichtsabfassung lässt sich folgendermaßen zusammenfassen:

- Strohsal, Till und Weber, Enzo (2013):
  Mean-Variance Cointegration and the Expectations Hypothesis
  Freie Universität Berlin, Berlin.
  Eine frühere Version ist als SFB 649 Discussion Paper 2011-007 erschienen sowie in Quantitative Finance im Erscheinen (http://dx.doi.org/10.1080/14697688.2013.814974).

  Eigenanteil: 50 Prozent

- Strohsal, Till und Winkelmann, Lars (2012):
  Assessing the Anchoring of Inflation Expectations
  Freie Universität Berlin, Berlin.
  Eine frühere Version ist als SFB 649 Discussion Paper 2012-022 erschienen.

  Eigenanteil: 50 Prozent

- Strohsal, Till (2013):
  Testing the Preferred-Habitat Theory: The Role of Time-Varying Risk Aversion
  Freie Universität Berlin, Berlin.

  Eigenanteil: 100 Prozent

- Strohsal, Till und Weber, Enzo (2013):
  The Signal of Volatility
  Freie Universität Berlin, Berlin.
  Eine frühere Version ist als SFB 649 Discussion Paper 2012-043 erschienen.

  Eigenanteil: 50 Prozent

Danksagung

Mein besonderer Dank gilt Dieter Nautz. Vor allem dafür, dass er sich so viel Zeit für die Betreuung dieser Arbeit genommen hat. Unzählige Anregungen und Ideen sowie viel konstruktive Kritik sind in diese Dissertation eingeflossen. Weiterhin möchte ich mich besonders bei Enzo Weber bedanken. Nicht nur die Koautorenschaft, sondern der fachlichen Austausch insgesamt waren eine enorme Bereicherung. Dies gilt ebenso für die unendlich wertvolle Unterstützung durch Jürgen Wolters. Für zahlreiche Diskussionen und Kommentare danke ich auch meinem Koautor Lars Winkelmann und meinem Zweitgutachter Helmut Lütkepohl. Schließlich möchte ich Jorien Korver dafür danken, mir mit viel Geduld die geschriebene Englische Sprache ein beachtliches Stück näher gebracht zu haben.

Der größte Dank gilt meiner Familie. Ich danke meiner Frau, meinen Eltern und meinen Schwiegereltern für die permanente Unterstützung und die Kraft, die sie mir dadurch gegeben haben.

Contents

General Introduction and Results

The research effort to understand the term structure of interest rates is vast. Despite a large number of competing theoretical models and even competing paradigms, there is general agreement on one aspect: The current shape of the yield curve contains information on expected future developments of macroeconomic variables (Gürkaynak and Wright 2012). For policy makers, it is central to extract this information and to influence expectations of market participants. Therefore, the term structure constitutes an integral part of monetary policy. But what kind of information do the yields comprise? And why is it so important?

The yield curve contains market expectations about future interest rates. This powerful result is based on the most widespread theory of interest rates, the expectations hypothesis. It states that longer-term rates represent simple averages of expected future short-term rates. If the yield curve is upward sloping, short-rates are expected to rise. Yet, the extent to what this prediction applies depends on the size of the risk premium which is demanded by investors to hold longer-term bonds. In fact, the literature documents a fairly poor empirical performance of the expectations hypothesis. The first part of this thesis shows, however, that some of the negative evidence can be rationalized. It is argued that, after all, the expectations hypothesis should be considered as a reasonable first point of reference when we think about the relation among interest rates of different maturities.

The yield curve also contains market expectations about future inflation rates. According to the Fisher hypothesis, the difference between nominal and real interest rates provides a measure of expected inflation. Over the last two decades, observations of real interest rates have become available to researchers. In addition to the US, many European countries such as France, Germany, the UK or Sweden have developed very large markets of inflation indexed bonds. Since these bonds adjust nominal payments by realized inflation, their yields can be considered as real interest rates. Together with nominal yields, a term structure of break-even inflation rates can be constructed (Gürkaynak et al. 2007, 2010a). Break-even rates indicate expected inflation and hence represent very useful information. According to the forward looking Phillips curve – a standard element of New Keynesian models – controlling inflation has a lot to do with managing expectations. Therefore, it is decisive for monetary policy whether inflation expectations are well anchored. Since it is not clear how the anchoring of expectations can be assessed empirically, the second part of this thesis proposes an answer to that question.

While there is no doubt among economists that the term structure contains valuable information for monetary policy, it remains debatable whether the yield curve can be fully understood if we restrict our attention to only one type of representative economic agent. According to the majority of term structure models, as the affine class, there solely exist risk-averse arbitrageurs who trade across the whole maturity spectrum. On the contrary, recent advances in the literature on preferred-habitat models emphasize that bond yields do not only reflect the expectations of arbitrageurs. A further type of agent, i.e. preferred-habitat investors with a preference for specific maturities, is introduced. These models predict that maturity-specific demand and supply affects bond prices (Greenwood and Vayanos 2010). For practitioners, this implies that changing the supply and hence the maturity structure of government debt can represent an alternative tool for monetary policy. The third part of this thesis is dedicated to an analysis of the statistical significance and the economic relevance of preferred-habitat effects.

In view of the methodology deployed in this thesis, linear and static approaches do often not fully meet the econometric demands of testable economic hypotheses. The predicted relations are frequently supposed to be time-varying. Market volatility plays a key role in governing this time variation. In case of a time-varying risk premium, as in the first paper, volatility can represent the amount of risk. In case of time-varying risk aversion, as in the third paper, volatility can proxy the sensitivity of economic agents to a given amount of risk. In both cases, volatility is connected to risk or uncertainty. This view is deeply rooted in the financial literature (Engle et al. 1987, Bali and Engle 2010). It exists, however, another prominent academic understanding which attributes volatility to the flow of information (Fleming et al. 1998, Gagnon and Karolyi 2009). The two concepts are not mutually exclusive but quite contrary. Exploring the nature of volatility should therefore provide important insights in what volatility actually measures and how we should exploit its time variation in empirical studies. The last part of this thesis addresses precisely these questions.

All four papers of the present thesis are devoted to contribute new evidence and to introduce new methodological approaches to the field of empirical macroeconomics. In the following the main contributions and results of each individual paper are briefly summarized.

- ***Paper 1***: *Mean-Variance Cointegration and the Expectations Hypothesis*

  This paper sheds further light on a well-known (alleged) violation of the expectations hypothesis of the term structure - the frequent finding of unit roots in interest rate spreads. It is shown that the expectations hypothesis implies that the non-stationarity stems from the holding premium, which is hence cointegrated with the spread. Within a stochastic discount factor framework the premium is modeled as being driven by the integrated variance of excess returns. A test for cointegration between mean and variance of random variables is introduced and applied to US bond data. The mean-variance cointegration test provides strong evidence for a long-run relation between conditional first and second moments. The findings suggest that the expectations hypothesis performs much better than we might have thought.

- ***Paper 2***: *Assessing the Anchoring of Inflation Expectations*

  This paper proposes a new approach to assess the degree of anchoring of inflation expectations. The implicit unit root assumption of the predominant news regressions is relaxed by introducing ESTAR non-linearities. The approach assumes globally stationary expectations and provides estimates of a market-perceived inflation target as well as the strength of the anchor that holds expectations at that target. A cross-country study is conducted based on a new data set of daily break-even inflation rates for the US, EMU, UK and Sweden. In contrast to the news regressions, which would have found unanchored expectations in all countries, the ESTAR results show that the degree of anchoring varies significantly across countries and horizons of expectations. While inflation expectations appear well anchored in the EMU, reasonably stable in the US and Sweden, they show signs of de-anchoring in the UK.

- ***Paper 3***: *Testing the Preferred-Habitat Theory: The Role of Time-Varying Risk Aversion*

  This paper tests the preferred-habitat model of Greenwood and Vayanos (2012). Special attention is paid to time-varying risk aversion which implies that the predicted positive relation between the term spread and relative supply of longer-term debt is stronger, when risk aversion is high. To capture this effect, a flexible time-varying coefficient model is introduced and applied to German bond data. The results provide supportive evidence for the model and indicate that the time variation is substantial: When risk aversion is high, yield spreads react by about 3 times as much as when risk aversion is low. The accumulated response of term spreads to a one standard deviation change in debt supply ranges between 5 and 33 basis points.

- ***Paper 4***: *The Signal of Volatility*

  This paper analyzes the inherently ambivalent economic interpretation of financial volatility in the academic literature. While volatility is considered an indicator of either information flow or uncertainty, it is shown in a stylized model economy that both views suggest volatility-dependent cross-market spillovers. Observing theses spillovers to decrease (increase) with volatility would favor the linkage to uncertainty (information). To this end, a simultaneous time-varying coefficient model is introduced in which structural ARCH-type variances serve two purposes: Governing the time variation of spillovers and ensuring statistical identification. Based on data of the US and further stock markets, part of the results support the connection between volatility and risk which was used in previous papers of this thesis.

Allgemeine Einführung und Ergebnisse

Die Forschungsanstrengungen zum Thema Zinsstruktur sind enorm. Trotz einer großen Anzahl an konkurrierenden theoretischen Modellen und widersprüchlichen empirischen Ergebnissen herrscht Einigkeit bezüglich eines Punktes: Die gegenwärtige Form der Zinsstrukturkurve enthält Informationen über die erwartete Entwicklung makroökonomischer Variablen (Gürkaynak and Wright 2012). Es ist ein zentrales Anliegen geldpolitischer Entscheidungsträger diese Informationen herauszufiltern und die Erwartungen der Marktteilnehmer zu beeinflussen. Die Zinsstruktur stellt daher einen integralen Bestandteil der Geldpolitik dar. Aber welche Informationen sind es genau, die in den Zinsen stecken? Und weshalb sind sie so bedeutend?

Zum einen enthält die Zinsstrukturkurve Informationen über erwartete Zinsen. Dieses fundamentale Ergebnis basiert auf der wohl am weitesten verbreiteten Theorie zur Zinsstruktur, der Erwartungshypothese. Die Theorie besagt, dass langfristige Zinsen schlicht Durchschnittswerte erwarteter kurzfristiger Zinsen darstellen. Entsprechend ist die Markterwartung, dass sich die kurzfristigen Zinsen erhöhen, wenn die Zinsstrukturkurve steigend ist. Die Gültigkeit dieser Aussage hängt jedoch maßgeblich von der Höhe der Risikoprämie ab, die Investoren verlangen, um langfristige Anleihen zu halten. Und so gilt die Erwartungshypothese, nach einhelliger Auffassung in der einschlägigen Literatur, empirisch in der Tat als sehr dürftig. Der erste Teil dieser Arbeit zeigt allerdings, dass ein Teil der empirischen Ergebnisse der Theorie nur vermeintlich widerspricht. Die Erwartungshypothese sollte dementsprechend als ein vernünftiger erster Referenzpunkt betrachtet werden, wenn über die Beziehungen zwischen Zinsen unterschiedlicher Fristigkeit nachgedacht wird.

Zum anderen enthält die Zinsstrukturkurve Informationen über erwartete Inflation. Nach dem Fisher-Effekt ist die Differenz zwischen nominalen und realen Zinsen ein Maß für Inflationserwartungen. Seit etwa zwei Jahrzehnten haben Wissenschaftler tatsächlich die Möglichkeit reale Zinsen zu beobachten. Neben den USA haben sich in mehreren europäischen Ländern wie etwa Frankreich, Deutschland, Großbritannien und Schweden große Märkte für inflationsindizierte Staatsanleihen entwickelt. Da diese Anleihen die nominalen Zahlungen an die Inflation anpassen, kann deren Verzinsung als real betrachtet werden. Zusammen mit den nominalen Zinsen ist es möglich, eine Strukturkurve sogenannter Break-Even-Inflationsraten zu berechnen (Gürkaynak et al. 2007, 2010a). Break-Even-Inflationsraten sind ein Maß für erwartete Inflation und beinhalten daher äußerst relevante Informationen. Gemäß der zukunftsgerichteten Phillipskurve - einem Standard-Element Neu-Keynesianischer Modelle - ist die Steuerung von Inflation eng verbunden mit der Steuerung von Inflationserwartungen. Es ist daher eine entscheidende Frage für die Geldpolitik, ob Inflationserwartungen gut verankert sind. Da es in der Literatur allerdings bisher unklar ist, wie diese Verankerung gemessen werden soll, schlägt das zweite Papier dieser Dissertation eine Antwort auf diese Frage vor.

Während es unter Ökonomen Konsens ist, dass Zinsen für die Geldpolitik wichtige Informationen enthalten, bleibt es strittig, ob ein theoretischer Analyserahmen, der nur einen repräsentativen Agenten berücksichtigt, zum Verstehen der Fristigkeitsstruktur ausreicht. Nach der Mehrzahl der Zinsstrukturmodelle, wie etwa der Affinen Modellklasse, existieren ausschließlich risikoaverse Arbitrageure, die über das gesamte Fristigkeitsspektrum hinweg Anleihen handeln. Im Gegensatz dazu betonen neuere theoretische Entwicklungen in der Literatur, dass Anleihezinsen nicht ausschließlich Erwartungen von Arbitrageuren widerspiegeln. Es wird ein weiterer Agententyp eingeführt: Investoren mit einer Präferenz für eine bestimmte Fristigkeit. Diese Modelle sagen vorher, dass Anleihezinsen auch durch laufzeitspezifische Angebote und Nachfragen beeinflusst werden (Greenwood and Vayanos 2010). Für die Praxis impliziert dies, dass An- oder Verkäufe von Anleihen, und damit die Beeinflussung der Fristigkeitsstruktur der Staatsverschuldung, ein alternatives Werkzeug für die Geldpolitik darstellen. Der dritte Teil dieser Arbeit analysiert die statistische und ökonomische Signifikanz lokaler Angebots- und Nachfrageeffekte.

Im Hinblick auf die in dieser Arbeit verwendete Methodik hat sich gezeigt, dass lineare statische Ansätze den ökonometrischen Anforderungen ökonomischer Hypothesen häufig nicht gerecht werden. So sind die theoretisch implizierten Beziehungen oftmals zeitvariierend. Marktvolatilität spielt bei der Beschreibung dieser Zeitvariation eine Schlüsselrolle. Im Fall einer zeitvariierenden Risikoprämie, wie im ersten Papier dargestellt, kann Volatilität für die Menge des Risikos stehen. Im Fall zeitvariierender Risikoaversion, wie im dritten Papier gezeigt wird, kann Volatilität die Sensibilität ökonomischer Akteure bezüglich einer gegebenen Menge an Risiko beschreiben. In beiden Fällen wird Volatilität mit Unsicherheit in Verbindung gebracht. Diese Ansicht ist in der Finanzierungsliteratur tief verwurzelt (Engle et al. 1987, Bali and Engle 2010). Daneben existiert jedoch ein weiteres akademisches Konzept, in dem Volatilität mit dem Fluss an Informationen gleichgesetzt wird (Fleming et al. 1998, Gagnon and Karolyi 2009). Die beiden Sichtweisen schließen sich nicht vollkommen gegenseitig aus, sind aber dennoch sehr gegensätzlich. Die Natur von Volatilität zu untersuchen, sollte daher wichtige Erkenntnisse darüber liefern, was durch Volatilität eigentlich gemessen wird und wie wir sie in empirischen Studien benutzen können. Der letzte Teil dieser Dissertation geht eben diesen Fragen nach.

Alle vier Papiere dieser Arbeit sollen dazu dienen, neue empirische Ergebnisse und methodische Ansätze zum Feld der empirischen Makroökonomie zu liefern. Im Folgenden sind die Beiträge und Resultate der einzelnen Aufsätze kurz zusammengefasst.

- ***Papier 1***: *Mean-Variance Cointegration and the Expectations Hypothesis*

  Dieses Papier untersucht eine bekannte (scheinbare) Verletzung der Erwartungshypothese der Zinsstruktur - das häufig auftretende Ergebnis einer Einheitswurzel in Zins-Spreads. Es wird gezeigt, dass die Erwartungshypothese impliziert, dass die Instationarität von der Risikoprämie stammt, welche daher mit dem Spread kointegriert. In einem Modellrahmen mit stochastischem Diskontierungsfaktor wird die Prämie durch die integrierte Varianz von Überschussrenditen beschrieben. Es wird ein Test auf Kointegration zwischen Mittelwert und Varianz von Zufallsvariablen eingeführt und auf Zinsdaten aus den USA angewandt. Der Mittelwert-Varianz Kointegrations Test liefert starke Hinweise für eine langfristige Beziehung zwischen bedingten ersten und zweiten Momenten. Die Ergebnisse deuten darauf hin, dass die Erwartungshypothese weitaus besser funktioniert als bisher angenommen.

- ***Papier 2***: *Assessing the Anchoring of Inflation Expectations*

  Dieses Papier schlägt einen neuen Ansatz zur Messung der Verankerung von Inflationserwartungen vor. Durch die Verwendung des ESTAR Modells wird die implizite Annahme einer Einheitswurzel im vorherrschenden Ansatz der News-Regressionen aufgegeben. Der ESTAR Ansatz nimmt stattdessen global-stationäre Erwartungen an. Er liefert Schätzungen eines vom Markt wahrgenommenen Inflationsziels sowie Schätzungen der Stärke des Ankers, der die Erwartungen an dem Inflationsziel hält. Die Arbeit beinhaltet eine länderübergreifende Studie, basierend auf einem neuen Datensatz täglicher Break-Even-Inflationsraten für die USA, EMU, UK und Schweden. Im Gegensatz zu den News-Regressionen, die in allen Ländern nicht verankerte Erwartungen gefunden hätten, zeigen die Ergebnisse der ESTAR Modelle, dass der Grad der Verankerung beträchtlich zwischen den Ländern und Erwartungshorizonten variiert. Während Inflationserwartungen in der EMU gut verankert und in den USA und Schweden recht stabil sind, erweisen sich Erwartungen in Großbritannien als nicht verankert.

- ***Papier 3***: *Testing the Preferred-Habitat Theory: The Role of Time-Varying Risk Aversion*

  Dieses Papier untersucht das Preferred-Habitat Modell von Greenwood und Vayanos (2012). Besondere Aufmerksamkeit wird der zeitabhängigen Risikoaversion zuteil, die besagt, dass die vom Modell implizierte positive Beziehung zwischen dem Zins-Spread und dem relativen Angebot an längerfristigen Schulden stärker ist, wenn die Risikoaversion hoch ist. Um diesen Effekt zu erfassen, wird ein flexibles ökonometrisches Modell mit zeitvariierenden Koeffizienten eingeführt und auf deutsche Anleihedaten angewandt. Die Ergebnisse stützen die ökonomischen Hypothesen und zeigen, dass die Zeitvariation von großer Bedeutung ist: Wenn die Risikoaversion hoch ist, reagieren Zins-Spreads etwa 3-mal so stark wie unter niedriger Risikoaversion. Die akkumulierte Reaktion von Zins-Spreads auf eine Änderung des relativen Angebots an Schulden in Höhe einer Standardabweichung liegt im Bereich zwischen 5 und 33 Basispunkten.

- ***Papier 4***: *The Signal of Volatility*

  Dieser Aufsatz analysiert die ambivalente ökonomische Interpretation von Finanzmarktvolatilität. In der Literatur wird Volatilität sowohl als Indikator für den Fluss an Informationen als auch für Unsicherheit betrachtet. Dieses Papier zeigt in einer stilisierten Modellökonomie, dass beide Ansichten Spillover-Effekte zwischen Märkten implizieren, die von der Volatilität abhängen. Eine Abnahme (Zunahme) dieser Spillover bei steigender Volatilität würde die Interpretation als Unsicherheit (Information) stützen. Um dies zu überprüfen, wird ein simultanes Modell mit zeitvariierenden Koeffizienten eingeführt, in dem strukturelle ARCH Varianzen zwei Zwecke erfüllen: Sie treiben die Zeitvariation der Spillover und garantieren statistische Identifikation. Basierend auf Daten aus den USA und weiteren Aktienmärkten, unterstützt ein Teil der Ergebnisse die Verwendung von Volatilität als Risikoindikator, so wie sie in anderen Teilen dieser Arbeit eingesetzt wurde.

1 Mean-Variance Cointegration and the Expectations Hypothesis

1.1 Introduction

The relation between interest rates of different maturities plays a key role in macroeconomics and finance. For monetary policy, the transmission mechanism from short to long rates is of particular importance. An obvious and plausible approach is given by the expectations hypothesis of the term structure (EHT), which remains one of the most examined as well as one of the most rejected theories.[1] The present paper focuses on the common implication of the EHT that interest rate spreads should be stationary, and provides an explanation why this property is almost never found in empirical studies. We show that this notorious lack of evidence can be attributed to a nonstationary term premium modeled by means of a stochastic discount factor model. Using our newly introduced *mean-variance cointegration* test, this explanation is verified by econometric results from unit root and cointegration analysis.

The implication of stationary spreads was first shown by Campbell and Shiller (1987). A popular linearized version of the EHT states that the spread equals expected future short rate changes plus a *constant term premium*, θ. Considering the two-period case, it is easy to see that for interest rates integrated of order one,

[1] A comprehensive survey covering early work and recent developments is provided by Gürkaynak and Wright (2012).

stationarity of the right-hand side in $Y_t^{(2)} - Y_t^{(1)} = 1/2\mathbb{E}[\Delta Y_{t+1}^{(1)}|I_t] + \theta$ goes hand in hand with cointegration on the left-hand side. Here, $Y_t^{(n)}$ denotes the yield on an n-period bond and $\mathbb{E}[\,\cdot\,|I_t]$ is the expectations operator, conditioning on the information available up to time t, I_t.

However, much evidence contradicts the implication of mean-reverting spreads. Among many others, Hall et al. (1992), Bremnes et al. (2001) and Hansen (2003) find that stationarity of the spreads is often not reflected in US data. The larger the difference in maturity the more often this outcome occurs. Wolters (1998) and Carstensen (2003) obtain the same result for German bond data. A number of authors argued that the assumption of a constant term premium may be unsuitable. Evidence for a time-varying premium is provided by Mankiw and Miron (1986), Evans and Lewis (1994), Hess and Kamara (2005), and Caporale and Caporale (2008), to name just a few. However, the term premium is unobservable and the EHT does not provide any guidance on how such a time-varying premium should be modeled.

Meanwhile, a great deal of literature has been produced that concerns the question of what exactly drives the commonly accepted time-variation in the term premium. One way of summing up the ongoing academic effort is to classify the different approaches within the broad class of stochastic discount factor (SDF) models. Detailed discussion of SDF models is provided by Cochrane (2001), and Balfoussia and Wickens (2007). Essentially, assets prices equal the expected discounted value of their future pay-offs. Yet, we emphasize that a time-varying but *stationary* premium that may be modeled by any particular SDF model does not change the EHT implication of stationary spreads.

The finding of nonstationary spreads is often interpreted as evidence against the validity of the EHT. This conclusion, however, ignores the possibility that the nonstationarity comes from the term premium, which is included in the interest rate spread. Therefore, the present paper argues that the unit root evidence can be reconciled with the expectations hypothesis if *integrated* spreads come along with *integrated* term premia (Hypothesis i). In that case, the nonstationarity puzzle would be rationalized if spreads and premia were *cointegrated*. This is what we label *mean-variance cointegration* (Hypothesis ii). Cointegration is required since, according to the expectations hypothesis, the difference between spread (mean) and premium (variance) leaves only stationary variables, namely first differences of the short rate.

In our analysis we apply the most simple, observable one-factor SDF model that is able to describe such an extremely persistent premium: the Sharpe-Lintner CAPM. The term premium is specified as the product of risk and its market price, equaling the expected excess return. The conditional second moment of excess returns thereby serves as the risk measure.[2] We estimate the term premium via a generalized autoregressive conditional heteroskedasticity (GARCH) model (Engle 1982, Bollerslev 1986) and show that the null of integrated conditional variance cannot be rejected. This result survives the inclusion of endogenous structural breaks under the alternative hypothesis: a form of nonlinearity that often induces artificial persistence. Finally, we propose a cointegration test and simulate the appropriate distribution of the test statistic. Empirically, we actually find cointegration relations between premia and spreads in US interest rate data. This explains the (seeming) violation of the necessary condition for the EHT to be valid, i.e., the frequent finding of nonstationary spreads.

The idea of cointegration between the conditional first and second moments can be extended to a number of other prominent topics in applied macroeconometrics. The Friedman (1977) hypothesis that inflation uncertainty has a negative effect on output growth and the hypothesis of Cukierman and Meltzer (1986) that inflation uncertainty increases the level of inflation imply interactions between means and variances of two different time series. Grier and Perry (2000), for instance, investigate these hypotheses in a bivariate setting (see also Fountas and Karanasos 2007). They fit ARMA-GARCH models to quarterly US output growth and inflation series. The mean equations are augmented by both conditional variance series as additional regressors that mostly appear insignificant. Using higher frequency inflation data and taking the possibility of mean-variance cointegration into account might show that the second moments actually have predictive power

[2]Common examples of that approach are Engle et al. (1987) and Bollerslev et al. (1988).

for inflation. Fountas et al. (2004) estimate univariate AR-GARCH-in-Mean models for quarterly Japanese real output growth data. According to unit root tests, their GDP growth series is stationary. Yet, their point estimates of the coefficients in the conditional variance equation suggest that the second moments may be integrated of order one. In that case, the regression with GDP growth rates would be unbalanced, suggesting a cointegration approach. This paper focuses on solving the nonstationarity puzzle of interest rate spreads. However, whenever conditional second moments appear in conditional mean equations, balancedness of the equation requires the two moments to have the same degree of integration. Our proposed cointegration test procedure can hence serve as a useful tool for other interesting applications in applied time series econometrics.

The paper proceeds as follows. Section 1.2 discusses stochastic discount factor models for term premia, looks at the EHT and derives two testable hypotheses. In section 1.3 we introduce the econometric methodology. In particular, we propose a procedure to test for mean-variance cointegration. This is followed by the presentation of the empirical results and several robustness checks. The final section provides a summary and contains concluding remarks.

1.2 Term Premium Models and the Expectations Hypothesis

In this section, firstly, the general framework of the SDF approach for modeling term premia is briefly outlined. This is followed by the presentation of the specific SDF model that we employ. Secondly, we turn to the relation between the term premium and the interest rate spread. Showing that the empirical finding of unit-root behavior in spreads can be explained by integrated term premia, we derive two testable hypotheses. Thirdly, it is illustrated how this explanation carries over from the SDF model to a (linearized) version of the EHT, which is prevalent in a large strand of literature.

1.2.1 The Stochastic Discount Factor Model and a CAPM-motivated Pricing Kernel

The SDF model relates the price of an asset to the expected present value of the future pay-off. Following the remarks of Smith and Wickens (2002) we have

$$P_t = \mathbb{E}[M_{t+1}X_{t+1}|I_t] \;, \tag{1.1}$$

where P_t denotes the price at time t. X_{t+1} represents the pay-off at $t+1$, M_{t+1} is the discount factor or pricing kernel ($0 \leq M_{t+1} \leq 1$) and $\mathbb{E}[\,\cdot\,|I_t]$ indicates the conditional expectation operator where the information set I_t contains all information available up to time t. As we are interested in the return $R_{t+1} = {}^{X_{t+1}}/_{P_t} - 1$, it is noted that

$$1 = \mathbb{E}[M_{t+1}(1 + R_{t+1})|I_t] \;. \tag{1.2}$$

By definition

$$\mathbb{E}[M_{t+1}(1+R_{t+1})|I_t] = \mathbb{E}[M_{t+1}|I_t]\mathbb{E}[1+R_{t+1}|I_t] + \mathrm{Cov}[M_{t+1},(1+R_{t+1})|I_t]$$

holds. Applying equation (1.2), the expected future gross return can be expressed as

$$\mathbb{E}[1+R_{t+1}|I_t] = \frac{1-\mathrm{Cov}[M_{t+1},(1+R_{t+1})|I_t]}{\mathbb{E}[M_{t+1}|I_t]} \;.$$

The return at $t+1$ from a riskless investment, denoted by r_t, is known at t and is hence included in the information set I_t. Therefore, regarding (1.2), this return produces the relation

$$\mathbb{E}[M_{t+1}|I_t] = \frac{1}{1+r_t} \;.$$

The latter equation allows us to write the *expected excess return* over the risk-free rate as

$$\mathbb{E}[R_{t+1}|I_t] - r_t = -(1+r_t)\,\mathrm{Cov}[M_{t+1},(1+R_{t+1})|I_t] \;. \tag{1.3}$$

Equation (1.3) represents the characteristic relation between risk and return. In SDF models, risk is measured as the covariance of the return with the variables that represent the discount factor M_{t+1}, in other words, the factors that enter the pricing kernel.

Smith and Wickens (2002) show in their survey that the SDF model can be seen as the umbrella framework that includes the most prominent asset pricing models. The SDF models proposed and investigated in the literature greatly differ in the specification of the discount factor. One possible classification relates to the nature of the factors as being either observable or latent variables.

In bond pricing, a recently widely used class is given by affine factor models. They assume the discount factor to be a linear function of the observable or unobservable factors. The Vasicek (1977) and the Cox, Ingersoll and Ross (1985) (CIR) models represent two of the most popular latent variable affine factor approaches. Dai and Singleton (2000) compare several multi-factor CIR models. In their influential study, Ang and Piazzesi (2003) augment a multi-factor Vasicek model by additional observable macroeconomic factors, thereby highlighting the importance of macroeconomic sources of risk for the short end, and that of latent factors for the long end of the term structure. Cochrane and Piazzesi (2005) show that one observable factor, a linear function of certain forward rates, can account for a huge part of the term premium. In recent literature on affine factor models, the intersection of macroeconomics and finance plays a prominent role; see Gürkaynak and Wright (2012) for a survey.

Moreover, there are two prime examples of implicit observable one-factor models: the CAPM (Sharpe 1964, Lintner 1965) and the CCAPM (Rubinstein 1976, Lucas 1978). Both models have a long tradition in finance, capturing the risk-return trade-off (see, e.g., Ghysels et al. 2005, Lundblad 2007 and Bali and Engle 2010). The CAPM represents the model of choice in the present paper. It implicitly assumes the factor to be the return on the market. The CAPM allows for an appealing economic interpretation due to the connection of risk as non-diversifiable return volatility. We will show that it fits the purpose of the underlying study well, i.e., explaining nonstationarity of spreads and introducing the concept of mean-variance cointegration. Combining this approach with more comprehensive risk models represents an attractive path for future research.

The CAPM can be classified as an implicit observable one-factor model and represents a very parsimonious choice. However, as will be seen below, it is well suited to account for the phenomenon of extremely persistent premia. The CAPM states that excess returns are described by

$$\mathbb{E}[R_{t+1} - r_t | I_t] = \lambda \cdot \mathrm{Cov}[R^m_{t+1}, R_{t+1} | I_t] \; . \tag{1.4}$$

In (1.4) R^m_{t+1} indicates the return on the market and, conditional on t,

$$\lambda = \frac{\mathbb{E}[R^m_{t+1} - r_t | I_t]}{\mathrm{Var}[R^m_{t+1} | I_t]}$$

is constant. Comparing the well-known equation (1.4) to (1.3), it becomes obvious that the CAPM may be understood as an implicit one-factor model with the discount factor

$$M_{t+1} = -\frac{\lambda}{1 + r_t}(1 + R^m_{t+1}) \; . \tag{1.5}$$

1.2.2 Stationarity Properties of Spreads and Premia: Testable Hypotheses

The present work is concerned with the term structure of interest rates and the explanation of nonstationary spreads. Hence, we shall proceed by deriving theoretical implications and by taking a closer look at the exact form of (1.4) in case of bond pricing and at the specific type of interest rate data that we investigate.

In order to focus primarily on the nonstationarity puzzle, we initially abstract from cross-asset and cross-market dependencies. We consider the most simple case of a stylized financial market that comprises only two assets: one risky and one riskless asset. The risky asset is represented by a coupon-carrying n-period bond with a yield to maturity (interest rate) of $Y^{(n)}_t$. The other asset is given by a one-period bond offering the riskless return $Y^{(1)}_t$. Let $H^{(n)}_{t+1}$ denote the return that one realizes at $t+1$ from holding the n-period bond for one period, i.e., from t to $t+1$

$$H^{(n)}_{t+1} = \frac{P^{(n-1)}_{t+1} - P^{(n)}_t + \; C}{P^{(n)}_t} \; . \tag{1.6}$$

Here, $P_t^{(n)}$ denotes the price that was paid at t and $P_{t+1}^{(n-1)}$ refers to the price of the bond at $t+1$, which now exists for one period and hence has only $n-1$ periods left until maturity. C represents the coupon payment. Since we later investigate holding returns constructed from yield data on bonds that are sold at par, we note that for these bonds by definition $P_t^{(n)} = 1$ and $C = Y_t^{(n)}$. The definition of excess holding returns of these data reduces to

$$H_{t+1}^{(n)} - Y_t^{(1)} = \underbrace{P_{t+1}^{(n-1)} - 1}_{c_{t+1}} + \underbrace{Y_t^{(n)} - Y_t^{(1)}}_{s_t} . \tag{1.7}$$

According to (1.7), *realized* excess returns over the risk-free rate from holding the n-period bond consist of two components. The first component equals the capital gain (loss) over the holding period, $c_{t+1} = P_{t+1}^{(n-1)} - P_t^{(n)}$. The second component constitutes the excess interest income, or, the spread $s_t = Y_t^{(n)} - Y_t^{(1)}$. *Expected* excess holding returns – also referred to as the *holding premium*[3] – are found by applying the conditional expectations operator to (1.7):

$$\mathbb{E}[H_{t+1}^{(n)} - Y_t^{(1)}|I_t] = \underbrace{\phi_{t+1}^{(n)}}_{I(0)\ /\ I(1)} = \underbrace{\mathbb{E}[c_{t+1}|I_t]}_{I(0)} + \underbrace{s_t}_{I(0)\ /\ I(1)} . \tag{1.8}$$

Usually, the econometrician cannot observe expectations. From the right-hand side of (1.4), however, we know that they can be described by the second moments of excess returns. In the two-asset case, the conditional covariance with the market becomes the conditional variance of the excess holding return of the n-period bond itself. Hence, plugging the definition of realized excess returns from (1.7) into the SDF model (1.4) yields

$$\underbrace{\phi_{t+1}^{(n)}}_{I(0)\ /\ I(1)} = \lambda \cdot \mathrm{Var}[c_{t+1} + s_t|I_t] = \lambda \cdot \underbrace{\mathrm{Var}[c_{t+1}|I_t]}_{I(0)\ /\ I(1)} , \tag{1.9}$$

the SDF-CAPM, where λ refers to the same proportionality factor as in (1.4). From (1.8) and (1.9) we draw two testable hypotheses:

[3]The literature sometimes confusingly uses "term premium" as an umbrella term for forward, holding and rollover premium. Since the following work requires the use of exact definitions, we apply those from the notes of Shiller (1990).

Hypothesis (i): Equal Degree of Integration

Given $\mathbb{E}[c_{t+1}|I_t] \sim I(0)$ and interest rate levels are integrated of order one, the spread s_t and the holding premium $\phi_{t+1}^{(n)}$ are either both stationary ($I(0)$) or both nonstationary ($I(1)$).

Hypothesis (ii): Mean-Variance Cointegration

If spread and holding premium are nonstationary they must be cointegrated. The cointegrating vector of s_t and $\mathrm{Var}[c_{t+1}|I_t]$ equals $(1,\ -\lambda)$.

Hypothesis (i) follows from (1.8). Assume $\mathbb{E}[c_{t+1}|I_t] \sim I(0)$.[4] In order for the equation to be balanced, the degrees of integration of spread and holding premium must be equal. Hypothesis (ii) follows from (1.9). According to the SDF-CAPM model, the holding premium equals $\lambda \cdot \mathrm{Var}[c_{t+1}|I_t]$ and hence it is the conditional second moment of excess returns that must be cointegrated with the conditional first moment of the spread. We emphasize that the interest rate spread, as the second component of excess holding returns, plays no part in the conditional variance, as it is included in the information set I_t. Thus, if the interest rate spread is integrated of order one, it must in fact be cointegrated with the conditional variance of the corresponding capital gain series. This is what we label *mean-variance cointegration*. The cointegrating vector of s_t and $\mathrm{Var}[c_{t+1}|I_t]$ is $(1,\ -\lambda)$. In this model λ may be interpreted as the market price of risk (PoR).

1.2.3 Linkage to the Linearized Expectations Hypothesis

The way in which Hypotheses (i) and (ii), derived from the SDF-CAPM model, carry over to the frequently used linearized version of the EHT is briefly outlined here. The reasoning that nonstationary spreads can be explained by nonstationary holding premia is shown to be consistent with the EHT.

[4]Theoretically, c_{t+1} can be considered as a series of price changes. Since prices of efficient markets normally behave like random walks or more general $I(1)$ processes, agents would expect c_{t+1} to be $I(0)$. Indeed, as will be seen later, this property of capital gains is found in the data.

The well-known form of the EHT is essentially only a linearization of stochastic equations that define returns (prices) in a financial market in the absence of arbitrage. Following the considerations of Shiller (1979), the holding return - expressed in terms of yields to maturity - can be linearized by means of a Taylor expansion of order one. The linearized holding return is then simply substituted for $H_{t+1}^{(n)}$ in the definition of the *holding premium* $\mathbb{E}[H_{t+1}^{(n)} - Y_t^{(1)}|I_t] = \phi_{t+1}^{(n)}$. The solution of the resulting first order difference equation yields the familiar expression that relates the interest rate spread to expected future short rate differences plus a *rollover premium* (for details see Appendix 1.A):

$$\underbrace{\overbrace{Y_t^{(n)} - Y_t^{(1)}}^{\equiv s_t}}_{I(1)} = \underbrace{\sum_{k=1}^{n-1} \omega'(k)\mathbb{E}[\Delta Y_{t+k}^{(1)}|I_t]}_{I(0)} + \underbrace{\overbrace{\theta_t}^{\text{Rollover Premium}}}_{I(1)} , \tag{1.10}$$

$$\theta_t = \sum_{k=0}^{n-1} \omega(k)\phi_{t+k+1}^{(n-k)} \tag{1.11}$$

and $\omega(k) = \gamma^{\,k} \frac{1-\gamma}{1-\gamma^{\,n}}$ respectively $\omega'(k) = \gamma^{\,k} \frac{1-\gamma^{\,n-k}}{1-\gamma^{\,n}}$ with $\gamma = 1/(1+\bar{Y})$, $0 < \gamma < 1$.

Equation (1.10), the linearized expectations model, was the theoretical starting point of numerous empirical investigations into the expectations hypothesis of the term structure. The conclusion that spreads should be stationary can directly be drawn from the above representation of the spread as a weighted average of expected future short rate changes in (1.10). Given that interest rate series are integrated of order one, agents would expect the changes in $\sum_{k=1}^{n-1} \omega'(k)\mathbb{E}[\Delta Y_{t+k}^{(1)}|I_t]$ to be $I(0)$. Furthermore, if the rollover premium θ_t is assumed to be stationary, the same holds for the spread (Campbell and Shiller 1991).

As can be seen from (1.11), the rollover premium θ_t can be written as a weighted sum of successive holding premia, in which the first summand equals $\phi_{t+1}^{(n)}$ from (1.9); see also Shiller (1990). Therefore, theoretically, the orders of integration of the two different kinds of premia, $\phi_{t+1}^{(n)}$ and θ_t, are equal. The conclusion drawn from (1.8) that nonstationary spreads can be explained by nonstationary holding premia is consistent with the linearized expectations model in (1.10) that would include an integrated rollover premium. Following the CAPM-motivated SDF model (1.9) allows us to derive an estimable specification for the premium, depending on the conditional variance of capital gains, $\mathrm{Var}[c_{t+1}|I_t]$. The results then carry over to the linearized expectations model in (1.10), which takes no independent stance on how θ_t might be measured.

1.3 Econometric Modeling

The methodology to be introduced follows three steps designed to empirically investigate Hypotheses (i) and (ii):

Equal Degree of Integration

(a) We determine the order of integration of interest rate spreads (conditional means) to obtain evidence of whether assuming stationary premia is appropriate.

(b) If spreads are $I(1)$, a test for integrated premia (conditional variances) will follow.

Mean-Variance Cointegration

(c) If our findings show that the premia are actually nonstationary[5] too, we will test for cointegration with the spreads and will estimate the proportionality coefficient as well as the adjustment speed.

Hence, at first, we discuss how to test for unit roots in the conditional mean of a time series (the spread) that potentially exhibits heteroskedasticity. Secondly, the same will be done with respect to nonstationarity of the conditional variance of a time series (the capital gain). Finally, we introduce the mean-variance cointegration approach to test for cointegration between spread and premium.

[5]The term stationarity always refers to *weak covariance stationarity*.

1.3.1 Testing for Integrated Interest Rate Spreads

Step (a) is conducted in this section. Whether there can be unit roots in interest rates is debatable due to the zero lower bound and some upper bound that applies under regular circumstances. Nonetheless, in limited samples the $I(1)$ property is often found to be empirically reasonable. Eventually, the conclusion of Campbell and Shiller (1987) that spreads must be stationary holds, irrespective of the persistence of the true data generating process (DGP) of interest rates.

Since the present data exhibits heteroskedasticity, a usual property of financial time series, we allow innovations to follow GARCH processes. However, unit root tests under conditional heteroskedasticity should be carried out with caution. With regard to the impact of neglected GARCH on the (augmented) Dickey-Fuller (ADF, Dickey and Fuller 1979) test, see, e.g., Kim and Schmidt (1993), Ling and McAleer (2003) and the literature they refer to. Due to the invariance principle, the ADF test proves to be asymptotically robust to covariance stationary GARCH errors. Small-sample properties were, however, conjectured to be affected in case of very persistent variance processes. Seo (1999), for instance, proposes a more powerful test[6]. The distribution in his test depends on a nuisance parameter, the relative weight $0 \leq \tau \leq 1$, and is bounded between the DF distribution ($\tau = 1$) and the standard normal ($\tau = 0$). We will double-check the standard ADF test results by the Seo test at a later stage. The well-known ADF test equation, or the mean equation in the Seo test, is given by

$$\Delta x_{t+1} = \delta + \psi x_t + \sum_{i=1}^{q} \delta_i \Delta x_{t+1-i} + u_{t+1} \ , \tag{1.12}$$

[6] The test from Seo (1999) uses the information arising from conditional heteroskedasticity by means of joint maximum likelihood estimation (MLE) of the autoregressive and the GARCH parameters. However, Charles and Darné (2008) find that for many practically relevant GARCH parameter values (i.e., for a sum of the ARCH and GARCH coefficients between 0.8 and 1 and for a GARCH parameter larger than the ARCH parameter) the DF test performs better than the Seo test with respect to power and size. Recent work from Kourogenis and Pittis (2008) explicitly analyzes integrated GARCH (IGARCH) innovations in the context of standard unit root tests. The DF test is included in their Monte Carlo simulations as the special case of uncorrelated innovations and appears to perform surprisingly well in the IGARCH case.

where q denotes the lag length and u_{t+1} is (possibly heteroskedastic) white noise. Under the null of a unit root, the lagged level in (1.12) has no effect on Δx_{t+1}. The test statistic is given by the t value of $\hat{\psi}$. It does not, however, follow a t-distribution.

1.3.2 Testing for Integrated Holding Premia

Spreads found to be $I(1)$ in unit root tests could only be consistent with EHT if the premium, which we model as the conditional variance of capital gains, was nonstationary (Hypothesis i, step b). To set up a suitable test procedure, the conditional mean of capital gains is specified as an AR(p_c) process with GARCH(1,1) errors $\epsilon_{c,\,t+1}$:

$$\begin{aligned} c_{t+1} &= a_c + \sum_{i=1}^{p_c} a_{c,i} c_{t+1-i} + \epsilon_{c,\,t+1} \ , \\ h_{c,\,t+1} &= \omega_c + \alpha_c \epsilon_{c,\,t}^2 + \beta_c h_{c,\,t} \ , \end{aligned} \tag{1.13}$$

where $\mathrm{Var}[c_{t+1}|I_t] = \mathbb{E}[\epsilon_{c,\,t+1}^2|\ I_t] \equiv h_{c\,,t+1}$. The parsimonious (I)GARCH(1,1) specification is known to capture variance dynamics of most financial time series fairly well. This is also true for the present data. The IGARCH(1,1) hypothesis (see Engle and Bollerslev 1986), in which the slope coefficients in the conditional variance equation sum up to one, is usually checked by likelihood ratio (LR) tests. Lumsdaine (1995), however, shows that the LR test within a Monte Carlo investigation is quite oversized in small samples. Busch (2005) proposes a robust LR test based on quasi-MLE (QMLE). His test statistic proves to be well behaved in small samples. The correction term $k = 0.5(\mathbb{E}[\xi_t^4] - 1)$ with $\xi_t = {}^{\epsilon_t}/\sqrt{h_t}$ is calculated under the alternative of a covariance stationary GARCH process. In that case the test statistic

$$LR_R = -\frac{2}{k}\big(l(\hat{\boldsymbol{\theta}}_r) - l(\hat{\boldsymbol{\theta}}_u)\big) \tag{1.14}$$

has actual size close to nominal size, even for skewed disturbances.[7] In (1.14), $\hat{\boldsymbol{\theta}}_r$ and $\hat{\boldsymbol{\theta}}_u$ are the respectively restricted and unrestricted QMLEs for the parameter

[7] For further details, see Busch (2005).

vector $\boldsymbol{\theta}$. However, since persistence in variance is a central question here, we additionally conducted a small-sample simulation experiment. That is, we simulated the respective distribution of LR_R under the null of a DGP according to (1.13) with parameter vector $\hat{\boldsymbol{\theta}}_r$ ($\alpha_c + \beta_c = 1$). The resulting critical values will be applied in addition to the χ^2 quantiles. Moreover, we will provide evidence from test variants, allowing for endogenous structural breaks under the alternative hypothesis.

1.3.3 Testing for Mean-Variance Cointegration

We continue by discussing the methodological approach to examine Hypothesis (ii): If spreads and holding premia are both nonstationary, they must be cointegrated (step c). Thus, a test is presented for cointegration between the mean of the spread series and the variance of the capital gain series.

While one might test for cointegration by checking the residuals from a static regression for stationarity, this approach is known to produce biased estimates and to lack efficiency. In order to overcome these problems, we proceed by using the dynamic cointegration test proposed by Stock (1987). Critical values for the case of observable regressors are found in Banerjee et al. (1998) and Ericsson and MacKinnon (2002). The test equation naturally follows from our approach in the previous subsection through augmentation of the ADF test equation (1.12) by the integrated variance series $h_{c,t+1}$ from (1.13) (i.e., under the restriction that $\alpha_c + \beta_c = 1$):

$$\Delta s_{t+1} = a + \rho s_t + \gamma h_{c,t+1} + \sum_{i=1}^{p} a_i \Delta s_{t+1-i} + \epsilon_{t+1} \ . \qquad (1.15)$$

Additionally, we control for GARCH effects in ϵ_{t+1}. Hence, (1.15) and the process for ϵ_{t+1}^2 are estimated simultaneously by (Q)ML. Relation (1.15) describes an ECM for the interest rate spread. Note that the capital gain variance-in-mean of (1.15) is conditional on the information available at t. In view of (1.8) and (1.9), that is exactly what follows from economic theory: cointegration between s_t and $h_{c,t+1} \equiv \text{Var}[c_{t+1}|I_t]$. To gain efficiency, lagged differences of $h_{c,t+1}$ are not included since they were found insignificant. The established reasoning when testing for cointegration in an error correction framework applies: In case of cointegration, the common nonstationary factor of the two variables cancels out, so that the linear combination $z_t \equiv (1,\ \gamma/\rho)(s_t,\ h_{c,t+1})'$ with $\rho < 0$ represents a stationary time series. If so, the relation in levels should contribute significantly to the explanation of Δs_{t+1}. On the contrary; under the null z_t is nonstationary and thus ρ is zero.

Turning towards the to be applied critical values, one might first think of those provided by Banerjee et al. (1998) or Ericsson and MacKinnon (2002) for the case of one exogenous variable. Yet, in contrast to usual cointegration testing in an error correction framework, the proposed test equation (1.15) contains an unobservable regressor, the IGARCH series $h_{c,t+1}$ (the capital gain variance) estimated in a preceding step. Hence, (1.15) can be seen as a quasi-IGARCH-in-Mean cointegration test since the variance that enters the mean equation is driven by the squared residuals $\epsilon^2_{c,t+1}$ that originate from a different mean equation, namely relation (1.13). Although $h_{c,t+1}$ is a martingale, it is known that the properties of such a series deviate from those of a random walk in many aspects (see Nelson 1990 and Kim and Linton 2011). Besides, the innovations ϵ_{t+1} in (1.15) themselves are highly heteroskedastic[8], as was found in specification tests. To the best of our knowledge, a theory on testing for cointegration between a conditional mean and an estimated conditional variance series has not yet been developed. In order to fill this gap, we derive the distribution of the test statistic (the t-value of $\hat{\rho}$) via simulation (see Appendix 1.B).

At this point, a further property of realized excess holding returns on par bonds is noteworthy. As can be seen from their definition in (1.7), excess returns on coupon-carrying bonds consist of two components: the capital gain c_{t+1} and the interest rate spread s_t. Excess returns are generally known to exhibit only a very slight autocorrelation. However, as follows from Evans and Lewis (1994), this empirical result may also occur in case spreads are nonstationary since variation of capital gains is usually very high, relative to s_t. The persistence of c_{t+1} and $c_{t+1} + s_t$ is then statistically indistinguishable. In other words, due to a very low

[8]For a discussion of potential consequences of GARCH effects on standard cointegration tests see, e.g., Seo (2007) and the literature cited therein.

signal-to-noise ratio, statistical tests fail to detect the true order of integration of excess returns in case of s_t being $I(1)$. This fact may also underlie the general difficulties of empirical finance to find a significant risk-return trade-off when trying to explain (statistically) strongly mean reverting excess returns by highly persistent second moments. Importantly, our theoretical result remains: the orders of integration of s_t and ϕ_{t+1} must be equal. Therefore, if there is cointegration, the ECM (1.15) estimates the PoR $\lambda = \frac{\gamma}{\rho}$ superconsistently and determines the long-run relation between the nonstationary component of excess returns, s_t, and the conditional variance of excess returns, $h_{c,t+1} \equiv \mathrm{Var}[H_{t+1}^{(n)} - Y_t^{(1)} | I_t] = \mathrm{Var}[c_{t+1} | I_t]$.

1.4 Empirical Results

1.4.1 Data

The subsequent analysis is based on weekly yields from 1/03/1992 to 12/29/2006 provided by the US Federal Reserve Statistical Release. The 15 years of US interest rate data should ensure a sufficient number of observations (783). All series are taken from the Treasury Constant Maturity data, which allows to directly compare these rates.[9] The sample period includes the timespan after the early 1990s recession and cuts off before the subprime and the euro crises. We choose this period to reduce the probability of breaks in the conditional first and second moments that often lead to artificial persistence, given that the true DGP exhibits such nonlinearities. As shown by Lamoureux and Lastrapes (1990a), structural breaks, which one would expect during the financial crisis, can produce spurious nonstationarity of a time series.[10]

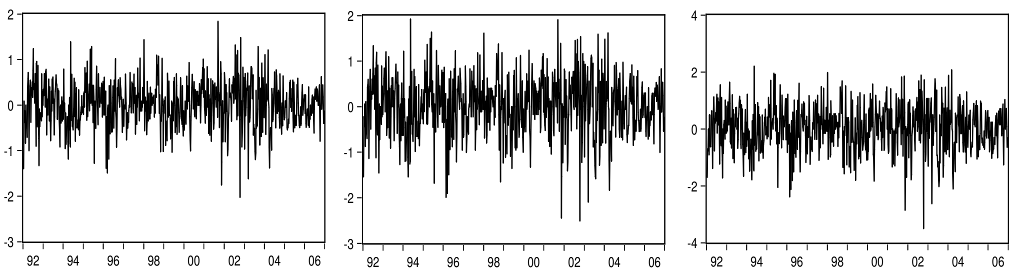
[9] All yields represent bond equivalent yields for securities that pay semi-annual interest.

[10] Tzavalis and Wickens (1995) argued that the regime of strong volatility of the very high interest rates in the early and mid-1980s was the cause of persistence in second moments of excess returns. In 1992, interest rates returned to their pre-early 1980s level and hence our sample should not be affected by that peak.

Excess holding returns and capital gains are calculated as, e.g., in Jones et al. (1998) or Christiansen (2000).[11] The calculation method is also described in Ibbotson and Associates (1994) and applied in Engle et al. (1987) for the case of effectively infinitely-lived bonds. Excess returns are defined as the return on holding a longer-term bond over one period in excess of the riskless rate. Longer-term bonds have maturities of 5, 7 and 10 years. The riskless rate is assumed to equal the standard 3-month Treasury rate, which from this point on will be referred to as the *short rate*. This is a common assumption and is considered to be the best alternative against using a one-week money market rate, which would imply, amongst other issues, discontinuities or outliers on settlement days (see Nelson 1991 or Jones et al. 1998)

As expected, mean excess holding returns increase with maturity of the long-term bond (0.032, 0.044 and 0.055). The same holds for the empirical standard deviations (0.488, 0.628 and 0.786). Capital gains, as part of the excess returns, are denoted by $C5$, $C7$ and $C10$. They equal the change in the present value (price) from one week to another (see Figure 1.1).

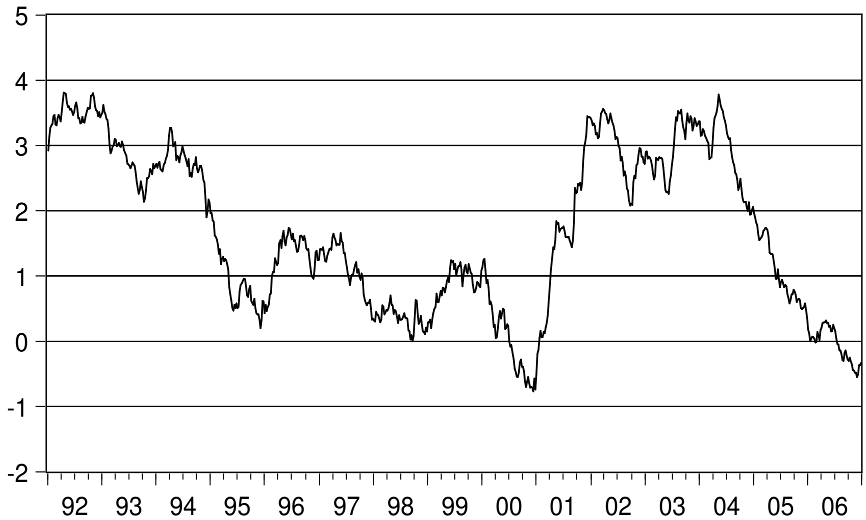
Figure 1.1: Weekly Capital Gain Series $C5$, $C7$ and $C10$ for 5-, 7- and 10-Year Bonds



[11] For the present yield data on par bonds and on a semi-annual basis, capital gains are defined as $c_{t+1} = P_{t+1}^{(n-1)} - 1 = \frac{1}{(1+\frac{1}{2}Y_{t+1}^{(n-1)})^{2n}} + \sum_{i=1}^{2n} \frac{\frac{1}{2}Y_t^{(n)}}{(1+\frac{1}{2}Y_{t+1}^{(n-1)})^j} - 1$; compare equations (1.6) and (1.7). The first term represents the present value of the principle and the second term represents those of the coupon payments. Since $Y_{t+1}^{(n-1)}$ is not available, $Y_{t+1}^{(n)}$ can be used instead. There should be no measurable difference between the yield of a 10-year bond and that of a bond with 9 years and 51 weeks to maturity, as pointed out by Shiller (1979), page 1197, footnote 8.

Since we use weekly observations of annualized interest rates, spreads are calculated as the fraction ($1/52$) of the difference between the respective long and short rate, that corresponds to a holding period of one week. Spread series are labeled $S5$, $S7$ and $S10$. By way of example, $S10$ is displayed in annualized form in Figure 1.2 ($S5$ and $S7$ exhibit a very similar shape).

Figure 1.2: Annualized Interest Rate Spread $S10$ between the 10-Year and the 3-Month Bond



1.4.2 Unit Root Tests for Interest Rate Spreads

In order to determine the integration order of interest rate spreads (Hypothesis i), the ADF test is applied. Note that, as far as levels are concerned, the test equation includes a constant, whereas the first differences do not have a deterministic part. A linear trend would not be meaningful for interest rate spreads and is also not supported by the data. The number of lagged differences is chosen according to the Schwarz information criterion (SIC). Hecq (1996) shows that standard information criteria can even be applied in the case of IGARCH and that the SIC performs best compared to the Hannan Quinn criterion (HQC) and the final prediction error (FPE). Since ADF test results are, however, known to be sensitive to the number of lagged differences in the test equation we double-checked our results by using HQC and FPE. Table 4.2 summarizes the unit root test results. It shows that nonstationarity is far from being rejected. Thus, all three spread series should be considered as integrated of order one.[12] When we apply HQC and FPE, both criteria suggest to include more lags but test statistics barely change. Performing Seo tests provides the same result: the null of a unit root cannot be rejected.[13]

Table 1.1: ADF Tests for Interest Rate Spreads

	Levels		First differences	
Variable	$\hat{t}$	q	$\hat{t}$	q
$S5$	−1.651 (0.456)	1	−22.916 (0.000)	0
$S7$	−1.466 (0.551)	1	−22.430 (0.000)	0
$S10$	−1.275 (0.643)	1	−22.021 (0.000)	0

Notes: Test statistics are denoted by $\hat{t}$. q refers to the number of lagged differences and p-values are given in parentheses.

1.4.3 IGARCH Tests for Holding Premia

So far we have seen that all spreads should be considered $I(1)$. As capital gains are clearly $I(0)$, following Hypothesis (i) (equal degree of integration), the EHT can only be valid in the presence of nonstationary holding premia. Since the

[12]The integration order of the interest rate series has been checked as well. According to ADF test results, there is very strong evidence that all interest rate series can be considered as integrated of order one.

[13]As the stationarity properties of interest rate spreads are crucial to the following analysis, we also conducted the KPSS test (Bartlett kernel and Newey-West bandwidth selection) with the null hypothesis of stationarity. This is to assure that non-rejections of nonstationarity are not simply due to the possible power problem of ADF-type unit root tests. For the present data, however, this does not seem to be the case, since the KPSS test clearly rejects stationarity. All test results that are not reported here can be obtained upon request.

SDF-CAPM model defines the holding premium as $\phi_{t+1}^{(n)} = \lambda \cdot \mathrm{Var}[c_{t+1}|I_t]$, testing Hypothesis (i) translates into testing for integrated capital gain variances; step (b).

We fit AR(p_c)-IGARCH(1,1) models to capital gain series, after which we test these models against the alternative hypothesis of autoregressive processes with covariance stationary variance series, that is, against AR(p_c)-GARCH(1,1) models. Table 1.2 summarizes the test results.

Table 1.2: Robust Likelihood Ratio Tests for Integrated Holding Premia

Variable	p_c	$\hat{\alpha}_c$	$\hat{\beta}_c$	$\hat{k}$	$\widehat{LR}_R$	$\chi^2_{0.90}(1)$
$C5$	1	0.035	0.951	1.598	2.648	2.706
		[0.014]	[0.022]			
$C7$	1	0.037	0.948	1.406	2.699	2.706
		[0.016]	[0.025]			
$C10$	1	0.033	0.953	1.241	2.163	2.706
		[0.015]	[0.024]			

Notes: Unrestricted model:

$$c_{t+1} = a_c + \sum_{i=1}^{p_c} a_{c,i} c_{t+1-i} + \epsilon_{c,\,t+1}\ ,$$
$$h_{c,\,t+1} = \omega_c + \alpha_c \epsilon^2_{c,\,t} + \beta_c h_{c,\,t}\ .$$

We tested for nonstationary holding premia $\phi_{t+1}^{(n)} = \lambda \cdot \mathrm{Var}[c_{t+1}|I_t]$, i.e., for integrated conditional variances, $\mathrm{Var}[c_{t+1}|I_t]$, of the respective capital gains on 5-, 7- and 10-year bonds, or $C5$, $C7$ and $C10$ respectively. $\hat{\alpha}_c$ and $\hat{\beta}_c$ refer to the unrestricted coefficient estimates. Under the null that $\alpha_c + \beta_c = 1$, the robust LR test statistic $LR_R = \frac{2}{k}\left(l(\hat{\boldsymbol{\theta}}_u) - l(\hat{\boldsymbol{\theta}}_r)\right)$ is $\chi^2(1)$. The estimated correction term is denoted by $\hat{k}$. ADF test statistics of -21.727, -13.363 and -13.290 for $C5$, $C7$ and $C10$ allow for a strong rejection of the null of nonstationarity.

Since the test statistic LR_R is $\chi^2(1)$ and $\chi^2_{0.90}(1) \approx 2.706$, the null of integrated variances cannot even be rejected at the 10% level. In all three cases we choose $p_c = 1$, following the SIC. Additionally, we conducted a small-sample experiment:

We simulated the distribution of LR_R for $C5$, $C7$ and $C10$ with DGPs under the null equal to our estimated AR(p_c)-IGARCH(1,1) models. For conditional normal distribution, a sample length of 783 observations and $100,000$ replications, the 90% quantiles turned out to be 3.418, 3.470 and 3.448. Thus, these results strengthened the test decision not to reject the null of IGARCH. It is, however, well known that spurious persistence can be caused by nonlinearities such as structural breaks neglected in the model specification (Lamoureux and Lastrapes 1990a). In our robustness section below, we also account for endogenous breaks in the unconditional variance.

1.4.4 Mean-Variance Cointegration Tests

So far, our results have shown that interest rate spreads can be treated as $I(1)$. The last section demonstrated that the three corresponding holding premia are also nonstationary. In order to test for Hypothesis (ii), cointegration between spreads and premia, we apply the mean-variance cointegration test introduced in section 1.3.3.

The following three ECMs were estimated:

$$\Delta S5_{t+1} = \underset{[-3.347]}{-0.0007} - \underset{[-3.967]}{0.022}\, S5_t + \underset{[4.047]}{0.005}\, \hat{h}_{c,\,t+1} + \underset{[5.685]}{0.192}\, \Delta S5_t + \hat{\epsilon}_{t+1}\ , \quad (1.16)$$

$$\Delta S7_{t+1} = \underset{[-2.821]}{-0.0006} - \underset{[-3.455]}{0.016}\, S7_t + \underset{[3.430]}{0.003}\, \hat{h}_{c,\,t+1} + \underset{[6.001]}{0.216}\, \Delta S7_t + \hat{\epsilon}_{t+1}\ , \quad (1.17)$$

$$\Delta S10_{t+1} = \underset{[-3.898]}{-0.0008} - \underset{[-4.116]}{0.017}\, S10_t + \underset{[4.435]}{0.002}\, \hat{h}_{c,\,t+1} + \underset{[6.025]}{0.222}\, \Delta S10_t + \hat{\epsilon}_{t+1}\ . \quad (1.18)$$

Compared to the test statistics from ADF tests in Table 4.2, t-values of the lagged level s_t increased considerably in (1.16), (1.17) and (1.18). Again, the lag length is chosen according to the SIC[14] and supported by specification tests for no residual

[14]Using different information criteria does not change cointegration test results.

autocorrelation. Furthermore, we allowed the residuals in (1.16), (1.17) and (1.18) to be GARCH(1,1) as well. Q statistics of standardized squared residuals, as well as LM tests for remaining GARCH, show that the parsimonious GARCH(1,1) specification proves to be reasonable.

Table 1.3 includes individual 1%, 5% and 10% simulated critical values for each of the three models.[15] There is only slight variation as the DGPs are very similar. Test results are unequivocal: In models (1.16) and (1.18) the null of no cointegration can be rejected at the 1% level. In (1.17) we reject the null at the 5% level. This is considered as strong evidence in favor of the existence of a cointegration relation. Economically, this means that a long-run equilibrium between US interest rate spreads and the corresponding one-period holding premia does in fact exist.

Table 1.3: Critical Values - Mean-Variance Cointegration Test

Model	$\hat{t}$	1%	5%	10%
$\Delta S5$	-3.967	-3.842	-3.247	-2.946
$\Delta S7$	-3.455	-3.838	-3.249	-2.939
$\Delta S10$	-4.116	-3.830	-3.247	-2.949

Notes: $\hat{t}$ refers to the estimated t value of the lagged level s_t in (1.16), (1.17) and (1.18).

Moreover, following our discussion in section 1.2, the coefficient $\lambda = \frac{\gamma}{\rho}$ in the respective attractor can be interpreted as the PoR. As mentioned earlier, we do not estimate the PoR via the standard GARCH-in-Mean model on the basis of excess return data. Instead, the PoR is estimated in a cointegration relation between the nonstationary component of the excess return, that is to say, the spread, and the integrated variance of the capital gain, meaning the component of excess returns not included in the conditioning information set. Thereby, (super)

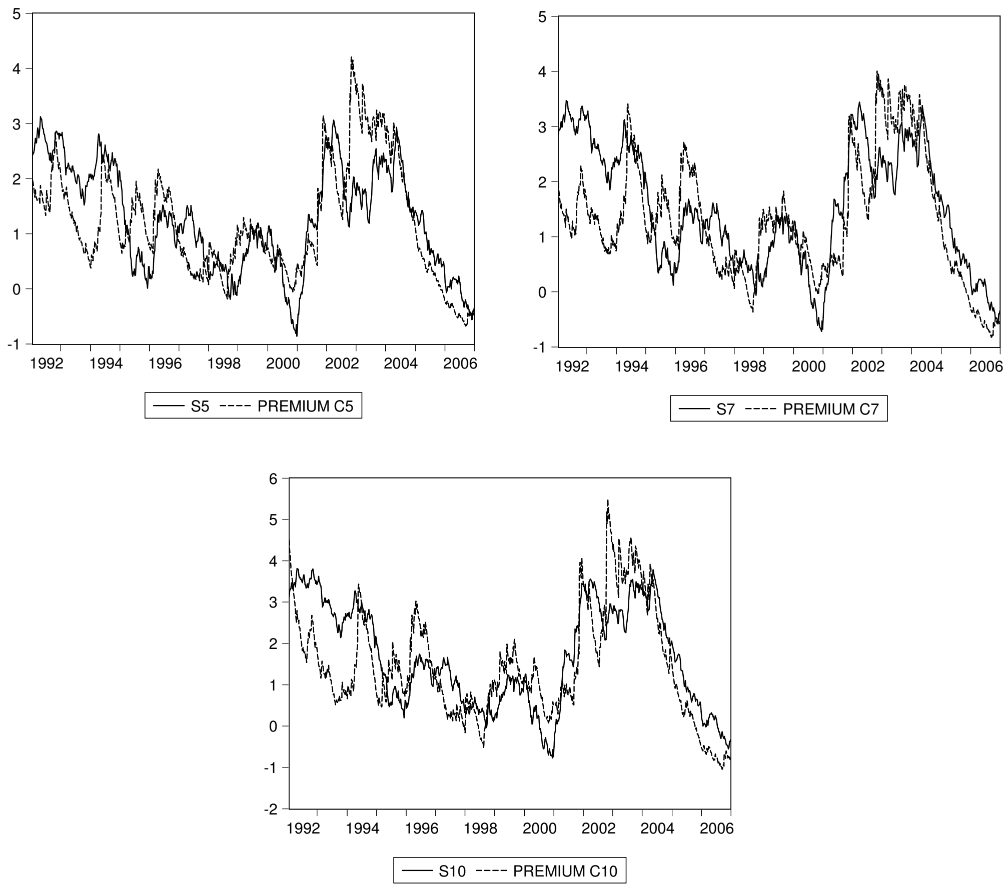
[15] As concerns the simulated critical values, some experiments made it clear that these depend on the conditional variance parameters of the DGPs of the capital gain variance and the spread series; see Appendix 1.B.

consistency of the estimator is guaranteed by the existence of cointegration. In (1.16), (1.17) and (1.18) we estimated PoRs of 0.228, 0.164 and 0.127. Compared to other findings, these coefficients are relatively but not implausibly small (see, e.g., Tzavalis and Wickens 1995, Bali and Engle 2010). For example, investors holding 7-year Treasury bonds expect at t that, on average, the excess return they will realize at $t+1$ will equal about one sixth of its variance. Assuming that the conditional variance at t would equal the (empirical) unconditional variance of 0.381, the excess return expected for $t+1$ becomes $0.164 \cdot 0.381 = 0.06\%$.

Regarding the above ECMs, adjustment coefficients may appear quite small at first glance. If, for instance, the spread $S10_t$ exceeds its equilibrium value by one unit (one percentage point), the spread will decrease by 0.017 units (percentage points) over the next week. However, after 13 weeks (one quarter) the initial equilibrium error has reduced to 0.771 and the half-life of the shock implied by system (1.18) equals just 33 weeks.

Figure 1.3 illustrates the long-run relation $s = \frac{a}{\rho} + \frac{\gamma}{\rho} h_c$ in the respective ECMs. A graphical analysis supports the statistical results: The general impression from Figure 1.3 is that there is quite a strong co-movement between the spreads and the corresponding one-period holding premia. From the beginning of the sample until the New Economy boom, the average level of all spreads decreases along with the volatilities. We see three noticeable peaks during that period: in 1994; between 1996 and 1997; and around the turn of the millennium. Whereas the first one results from long rates rising faster than the short rate, the second and third peaks were triggered by increasing long rates when the short rate remained roughly constant. In view of equation (1.8), rising spreads are associated with growing holding premia and hence with rising capital gain variances. Indeed, we clearly see that the variance movement features similar peaks. However, the period after 2001 when the short rate fell steeply for several years and, accordingly, spreads went up, is most striking. In support of the cointegration test results, this timespan is also characterized by high volatility.

Figure 1.3: Spreads and Corresponding One-Period Holding Premia



1.5 Robustness Checks

Finally, we conducted several robustness checks. Among other things, we were concerned with three main issues: spurious persistence in the conditional variance due to neglected nonlinearities such as structural breaks; initial value issues; and the conditional normal distribution assumption.

1.5.1 Endogenous Structural Breaks and Persistence in Variance

We would like to stress the point that persistence in conditional second moments can be a result of neglected structural change in the variance. Lamoureux and Lastrapes (1990a) provide examples of that phenomenon. A specific example is also found in Tzavalis and Wickens (1995), who show that the persistence in volatility of US holding premia between 1970 and 1986 is the result of a structural shift during a period of exceptionally high variances (October 1979 - September 1982). In general, the timing of structural breaks is quite difficult. In order to avoid the arbitrariness of choosing break dates exogenously, we conducted an endogenous break search. We therefore augmented the unrestricted conditional variance equation by a shift dummy and selected the date where the dummy had the highest t-value. As shown in the unit root literature (e.g. Zivot and Andrews 1992), the additional step of estimating the break date affects the distribution of the test statistic. We therefore simulated the distribution, allowing for a break in the GARCH constant under the alternative hypothesis. The results show that none of the three capital gain series allow a rejection of the null of IGARCH at the 10% level. We also allowed two level shifts with endogenous break dates. This improved the likelihood under the alternative hypothesis only slightly, so that nonstationarity was not rejected in this case, either.

1.5.2 Initial Values and the Shape of the Variance Series

The choice of initial values has no impact on our simulation results. However, the shape of the estimated capital gain variance series varies slightly, particularly during the first year (about 52 observations) of initializing GARCH models using, for instance, backcast exponential smoothing (where $h_0 = {}^{\kappa^N}\!/_N \sum_{t=0}^{N} \hat{\epsilon}_t^2 + (1-\kappa) \sum_{j=0}^{N} \kappa^{N-j-1} \hat{\epsilon}_{N-j}^2$, $0 < \kappa \leq 1$) instead of simply the mean of squares of residuals. Since the initial value impact essentially vanished after about one year, we re-estimated the ECMs starting the sample at the beginning of 1993 using 731 observations.[16] Compared to the estimates in (1.16), (1.17) and (1.18), test statistics decreased (i.e., increased in absolute value) to -4.512, -4.476 and -4.402. The lower number of observations affects critical values only at the second decimal place. We therefore reject the null of no cointegration at the 1% level in all ECMs.

1.5.3 Distributional Assumption

We now turn to investigating how the distributional assumption affects our test and simulation results. First of all, we drew random samples for $\xi_{c,\,t+1} = {}^{\epsilon_{c,\,t+1}}\!/\!\sqrt{h_{c,\,t+1}}$ and $\xi_{s,\,t+1} = {}^{\epsilon_{s,\,t+1}}\!/\!\sqrt{h_{s,\,t+1}}$ from Student's t-distributions in both simulation experiments: the LR test for IGARCH (section 1.3.2) and the mean-variance cointegration test (section 1.3.3). Degrees of freedom are set to be equal to estimated values under the null and lie between 8 and 18. Most of the estimated values are clearly larger than 10, indicating that the initial assumption of normality is not violated too strongly for the present data. Since sample excess kurtosis of all series is relatively small (between 0.8 and 1.5), this is not surprising. In order to analyze the effect of possibly incorrectly specified innovations our estimates actually become QMLEs in the sense that Gaussian likelihoods are maximized even though we have generated t-innovations. As expected, the smaller the number of degrees of freedom the more the distribution of LR_R shifts to the right. Hence, the decision not to reject the null of integrated variances in section 1.4.3 is strengthened. Similarly, the distribution of the t-value of $\hat{\rho}$ also shifts to the right, so that we can reject the null of no cointegration in section 1.4.4 at an even higher significance level (since the t-values are negative).

1.6 Conclusion

The present paper empirically examines a well-known implication of the expectations hypothesis of the term structure (EHT), namely, interest rate spreads

[16]Persistence in variances proves not to be affected by initial conditions.

should be stationary. We shed more light on the question why there is much evidence that contradicts the implication of mean-reverting spreads; see Hall et al. (1992) or Hansen (2003).[17] This implication has also been the pivotal element in many studies that analyze the spread's predictive power for short rate changes or other macroeconomic variables such as inflation and GDP growth (Mankiw and Miron 1986, Kugler 1988 or Caporale and Caporale 2008). The consequences of theoretically implied stationarity properties of interest rate spreads are obviously extensive. We are therefore concerned with the question why they are almost never met and argue that nonstationary holding premia can provide an answer.

The theoretical starting point is the one-period holding premium, defined as the sum of interest rate spread and expected capital gain, from which two testable hypotheses are derived. Hypothesis (i): Given stationary capital gains, spread and holding premium must exhibit the same order of integration. Hypothesis (ii): If this order equals one, spread and holding premium must be cointegrated. With respect to the economic and econometric modeling of spread and premium, we refer to mean-variance cointegration.

We show that explaining a nonstationary spread by integrated holding premia is consistent with the frequently used linearized version of the EHT. The latter would include a nonstationary rollover premium that we explicitly link to the holding premium. In modeling and estimating the holding premium, we employ an observable single-factor SDF model with a CAPM-motivated pricing kernel. The holding premium is proportional to the conditional variance of excess returns. In order to test for Hypothesis (i), unit root tests are applied for spreads and robust LR tests for IGARCH variances of excess returns. To test for Hypothesis (ii), a mean-variance cointegration test in an error correction framework is proposed and the small-sample distribution of the test statistic is derived through simulation. Our approach may be seen as a quasi-IGARCH-in-Mean cointegration test as the variance that enters the mean equation is estimated in a preceding step and is driven by the squared residuals from a different mean equation.

[17] Further cointegration studies such as Shea (1992), Zhang (1993), Engsted and Tanggaard (1994), Johnson (1994), Wolters (1995), Pagan et al. (1996), Wolters (1998), Carstensen (2003) make the empirical finding of unit roots in interest rate spreads an almost stylized fact.

The empirical analysis is based on weekly observations of US Treasury Constant Maturity data. We examine three different spreads between the short rate (3-month Treasury rate) and long rates with maturities of 5, 7 and 10 years. Following the ADF test results, all spreads should be considered nonstationary. Further unit root tests unanimously confirm the nonstationarity of the spreads. Subsequently, estimating conditional variances of excess returns shows that the null hypothesis of IGARCH cannot be rejected. Additionally, this result holds when endogenous structural breaks are incorporated. Hence, we conclude that holding premia are also integrated. The most important step follows: Testing for cointegration between premia and spreads. As the main result of the present work, we actually find highly significant long-run relations between all spreads and corresponding premia.

Following the idea of arbitrage-free financial markets and rational expectations, the EHT provides a simple and appealing description of the relation between interest rates of different maturities. Long rates embody information on expected future short rates and both rates are tied together within a long-term equilibrium relation. This equilibrium can be captured by a cointegration relation. However, the modeling of the term premium plays a key role here. This third variable should be modeled carefully and sometimes, as in the present case, even be included in the cointegration relation. The present paper has shown that nonstationary spreads can be reconciled with the EHT if spreads and premia are cointegrated, i.e., in case of mean-variance cointegration. The basic statement of the EHT concerning the relation between interest rates of different maturities remains applicable when the premium is modeled by means of approaches from finance theory. While the present paper mainly focuses on the aspect of cointegration, it also underlines the relevance of the vast and still evolving literature on identifying economically interpretable driving forces of the premium (as e.g. Ang and Piazzesi 2003 or Gürkaynak and Wight 2012).

Several extensions of the present approach appear interesting. First, our approach could be generalized to non-diversifiable risk (e.g. Bollerslev et al. 1988 or Balfoussia and Wickens 2007). The frequent failure of the EHT would be explained by integrated covariance series that could be obtained from multivariate GARCH models. This would allow to control for cross-asset and cross-market dependencies. Second, since appropriate modeling of the persistence of the premium proved to be crucial, a further possible extension would be to allow for fractional integration in interest rates (Connolly et al. 2007) and conditional variances (Baillie et al. 1996). If, for example, the order of integration of spreads and premia is equal, but appears to be less than one, cointegration tests may be carried out in a fractionally (co)integrated framework. Third, as shown in the introduction, the mean-variance cointegration approach also bears significant potential for application in other macroeconometric topics. The empirical studies of Grier and Perry (2000) and Fountas et al. (2004) on the interactions of inflation (uncertainty) and output growth (uncertainty), inspired by the hypotheses of Friedman (1977) and Cuckierman and Meltzer (1986), are two examples of these topics. We, however, leave these issues for future research. [18]

[18] A slightly different version of this paper is forthcoming as Strohsal and Weber (2013). Part of the present work is based on results from my diploma thesis Strohsal (2009).

1.A The Linkage from the Holding Premium in the SDF-CAPM Model to the Rollover Premium in the Linearized Expectations Model

The conclusion drawn from the SDF-CAPM model that nonstationary spreads can be explained by nonstationary *holding premia* is consistent with the familiar linearized version of the EHT that would include an integrated *rollover premium.* The well-known form of the EHT that will be derived now is essentially a linearization of equations that define returns (prices) in a financial market in the absence of arbitrage.

Firstly, consider the definition of the yield to maturity of an n-period bond

$$P_t^{(n)} = \frac{C}{(1+Y_t^{(n)})} + \frac{C}{(1+Y_t^{(n)})^2} + \cdots + \frac{1+C}{(1+Y_t^{(n)})^n} \,.$$

Most compactly, this can be written as:

$$P_t^{(n)} = \frac{C}{Y_t^{(n)}} + \frac{Y_t^{(n)} - C}{Y_t^{(n)}(1+Y_t^{(n)})^n} \tag{1.19}$$

The one-period holding return defined in (1.6) can be expressed in terms of yields to maturity by using (1.19), so that

$$H_{t+1}^{(n)} = \frac{\frac{C}{Y_{t+1}^{(n-1)}} + \frac{Y_{t+1}^{(n-1)} - C}{Y_{t+1}^{(n-1)}(1+Y_{t+1}^{(n-1)})^{n-1}} + C}{\frac{C}{Y_t^{(n)}} + \frac{Y_t^{(n)} - C}{Y_t^{(n)}(1+Y_t^{(n)})^n}} - 1\,. \tag{1.20}$$

According to the considerations of Shiller (1979) we linearize (1.20) via a Taylor expansion of order one. When $H_{t+1}^{(n)}(Y_t^{(n)},\ Y_{t+1}^{(n-1)},\ C)$ is considered a function of three variables, we know from Taylor's theorem that in the neighborhood of $Y_t^{(n)} = Y_{t+1}^{(n-1)} = C = \bar{Y}$ it holds that

$$
\begin{aligned}
H_{t+1}^{(n)} \approx H_{t+1}^{(n)}(\bar{Y},\ \bar{Y},\ \bar{Y}) + & \frac{\partial H_{t+1}^{(n)}}{\partial Y_t^{(n)}}|_{Y_t^{(n)}=Y_{t+1}^{(n-1)}=C=\bar{Y}} \cdot (Y_t^{(n)} - \bar{Y}) \\
+ & \frac{\partial H_{t+1}^{(n)}}{\partial Y_{t+1}^{(n-1)}}|_{Y_t^{(n)}=Y_{t+1}^{(n-1)}=C=\bar{Y}} \cdot (Y_{t+1}^{(n-1)} - \bar{Y}) \\
+ & \frac{\partial H_{t+1}^{(n)}}{\partial C}|_{Y_t^{(n)}=Y_{t+1}^{(n-1)}=C=\bar{Y}} \cdot (C - \bar{Y}) \equiv H_{t+1}^{'(n)} .
\end{aligned}
$$

Plugging in (1.20) and evaluating the derivatives finally yields

$$
H_{t+1}^{'(n)} = \delta_n Y_t^{(n)} - (\delta_n - 1) Y_{t+1}^{(n-1)} , \tag{1.20'}
$$

where $\delta_n = 1 + \bar{Y}^{-1} - (\bar{Y}(1+\bar{Y})^{n-1})^{-1}$. If one applies $H_{t+1}^{'(n)} \approx H_{t+1}^{(n)}$ to the definition of the holding premium, $\mathbb{E}[H_{t+1}^{(n)} - Y_t^{(1)}|I_t] = \phi_{t+1}^{(n)}$, the resulting first order difference equation with variable coefficients is

$$
Y_t^{(n)} = \gamma_n \mathbb{E}[Y_{t+1}^{(n-1)}|I_t] + (1-\gamma_n)(\phi_{t+1}^{(n)} + Y_t^{(1)}) , \tag{1.21}
$$

where $\gamma_n = {}^{(\delta_n - 1)}/_{\delta_n}$. The solution of (1.21) can be derived by recursive substitution. Therefore, we initially use the law of iterated expectations and note that

$$
\begin{aligned}
\mathbb{E}[Y_{t+1}^{(n-1)}|I_t] =& \gamma_{n-1}\mathbb{E}[Y_{t+2}^{(n-2)}|I_t] + (1-\gamma_{n-1})\mathbb{E}[\phi_{t+2}^{(n-1)} + Y_{t+1}^{(1)}|I_t] \\
\mathbb{E}[Y_{t+2}^{(n-2)}|I_t] =& \gamma_{n-2}\mathbb{E}[Y_{t+3}^{(n-3)}|I_t] + (1-\gamma_{n-2})\mathbb{E}[\phi_{t+3}^{(n-2)} + Y_{t+2}^{(1)}|I_t] \\
& \vdots \\
\mathbb{E}[Y_{t+n-1}^{(1)}|I_t] =& \gamma_1\mathbb{E}[Y_{t+n}^{(0)}|I_t] + (1-\gamma_1)\mathbb{E}[\phi_{t+n}^{(1)} + Y_{t+n}^{(1)}|I_t] .
\end{aligned}
$$

Recursive substitution of the above expressions into (1.21) yields

$$
\begin{aligned}
Y_t^{(n)} = \overbrace{\gamma_n \cdot \gamma_{n-1} \cdot \ldots \cdot \gamma_1 Y_{t+n}^{(0)}}^{=0} &+ (1-\gamma_n)\mathbb{E}[\phi_{t+1}^{(n)} + Y_t^{(1)} | I_t] \\
&+ \gamma_n \cdot (1-\gamma_{n-1})\mathbb{E}[\phi_{t+2}^{(n-1)} + Y_{t+1}^{(1)} | I_t] \\
&+ \gamma_n \cdot \gamma_{n-1} \cdot (1-\gamma_{n-2})\mathbb{E}[\phi_{t+3}^{(n-2)} + Y_{t+2}^{(1)} | I_t] \\
&\vdots \\
&+ \gamma_n \cdot \gamma_{n-1} \cdot \gamma_{n-2} \cdot \ldots \cdot \gamma_2 (1-\gamma_1)\mathbb{E}[\phi_{t+n}^{(1)} + Y_{t+n}^{(1)} | I_t] \ .
\end{aligned}
$$

This can be shortened to

$$
Y_t^{(n)} = \sum_{k=0}^{n-1} \omega(k)\mathbb{E}[Y_{t+k}^{(1)} | I_t] + \sum_{k=0}^{n-1} \omega(k)\phi_{t+k+1}^{(n-k)} \ ,
$$

with the weighting scheme $\omega(k) = \gamma^{\,k} \frac{1-\gamma}{1-\gamma^{\,n}}$. Via the identity $Y_{t+k}^{(1)} = Y_t^{(1)} + \sum_{i=1}^{k} \Delta Y_{t+i}^{(1)}$ we obtain

$$
Y_t^{(n)} - Y_t^{(1)} = \sum_{k=1}^{n-1} \omega(k) \sum_{i=1}^{k} \mathbb{E}[\Delta Y_{t+i}^{(1)} | I_t] + \sum_{k=0}^{n-1} \omega(k)\phi_{t+k+1}^{(n-k)} \ .
$$

Rearranging terms produces the well-known expression for the interest rate spread, i.e.,,

$$
\overbrace{Y_t^{(n)} - Y_t^{(1)}}^{\equiv s_t^{(n)}} = \sum_{k=1}^{n-1} \omega'(k)\mathbb{E}[\Delta Y_{t+k}^{(1)} | I_t] \quad + \quad \overbrace{\theta_t}^{\text{Rollover Premium}} \quad , \tag{1.10}
$$

where $\omega'(k) = \gamma^{\,k} \frac{1-\gamma^{\,n-k}}{1-\gamma^{\,n}}$ and $\theta_t = \sum_{k=0}^{n-1} \omega(k)\phi_{t+k+1}^{(n-k)}$ with $\gamma = {}^{1}/_{(1+\bar{Y})}$, $0 < \gamma < 1$. This shows (1.10) and (1.11).

1.B The Mean-Variance Cointegration Test: Simulating the Distribution of the Test Statistic

As concerns the simulated critical values, some experiments made it clear that these depend on the parameters of the DGPs of the capital gain variance and the spread series. Under the null, these are the equations in (1.13) with $\alpha_c + \beta_c = 1$ and

$$\begin{aligned} \Delta s_{t+1} &= \sum_{i=1}^{p_s} a_{s,i} \Delta s_{t+1-i} + \epsilon_{s,\,t+1} \ , \\ h_{s,\,t+1} &= \omega_s + \alpha_s \epsilon_{s,\,t}^2 + \beta_s h_{s,\,t} \ , \end{aligned} \tag{1.22}$$

where the first-difference autoregression is implied by the unit root in the level of the spread series; $\mathrm{Var}[\Delta s_{t+1}|I_t] = \mathbb{E}[\epsilon_{s,\,t+1}^2|\ I_t] \equiv h_{s\,,t+1}$. Hence, we simulated three different distributions of the test statistic in the error correction models (1.16), (1.17) and (1.18) with parameters in (1.13) and (1.22) according to our empirical estimates. Figure 1.4 illustrates the dependence on the parameters of the variance process $h_{c,t+1}$; the DGP (1.13).

It contains two Epanechnikov kernel density estimates[19] of the distributions of the test statistics for $\alpha_c = 0.033$ and $\alpha_c = 0.3$ ($\beta_c = 1 - \alpha_c$) with all other parameters unchanged and equal to our estimates for $C10$ and $\Delta S10$ in (1.13) and (1.22). It can be seen that the increase in α_c shifts the distribution to the right. The mean (variance) changes from -1.841 (0.791) to -1.771 (0.765). Both distributions are slightly skewed (0.244 and 0.226) and exhibit kurtosis of 3.404 and 3.359. For $\alpha_c \to 0$ the distribution moves leftwards but critical values change only at the second decimal place. The subsequent steps sketch the simulation of the test statistic of the mean-variance cointegration test proposed in section 1.3.3.

Step 1. Set initial values $h_{c,\,0}$ and $h_{s,\,0}$ in (1.13) and (1.22) equal to the mean of squares of $\hat{\epsilon}_{c,\,t+1}$ and $\hat{\epsilon}_{s,\,t+1}$, respectively.[20]

Step 2. Draw two random samples of size $N = 783$ (equal to the number of observations in the present analysis) from a standard normal distribution.

[19] We use a data-based bandwidth selection according to Silverman (1986).

[20] Dependence on the initial values turned out to be negligible. The initial values of s_t, $t = 0, ..., p_s$ are arbitrarily chosen in the sense that they are realizations of two standard normally distributed random variables.

These random shocks are denoted by $\xi_{c,\,t+1}$ and $\xi_{s,\,t+1}$.

Step 3. Generate data recursively according to (1.13) and (1.22) with $\epsilon_{c,\,t+1} = \xi_{c,\,t+1}\sqrt{h_{c,\,t+1}}$ and $\epsilon_{s,\,t+1} = \xi_{s,\,t+1}\sqrt{h_{s,\,t+1}}$.

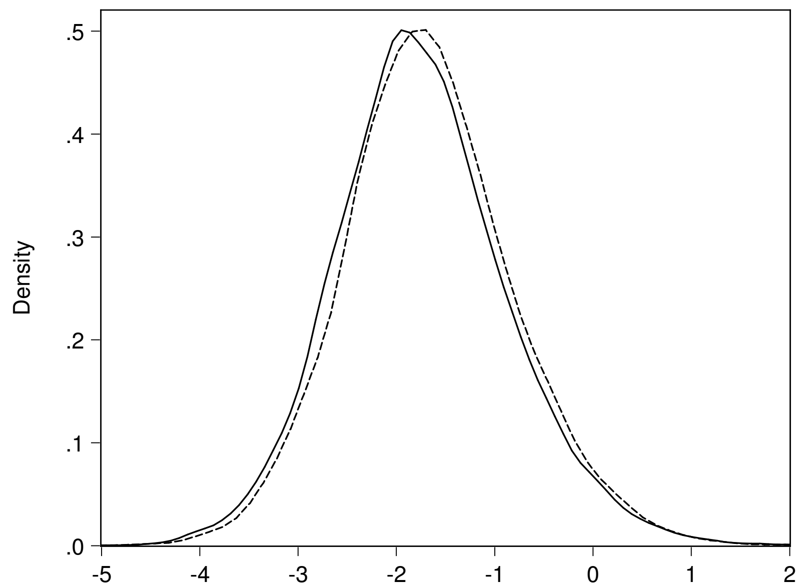
Step 4. Estimate model (1.13) via ML (BHHH algorithm) and save $\hat{h}_{c,\,t+1}$.

Step 5. Estimate model (1.15) via ML, using the generated spread series from *Step 3* and the estimated capital gain variance from *Step 4*, and save the t-value of $\hat{\rho}$ based on robust standard errors following Bollerslev and Wooldridge (1992).

Step 6. Repeat *Step 1* to *Step 5* 100,000 times.

Step 7. Calculate the 1.00, 5.00 and 10.00 percentiles from the distribution of the t-value of $\hat{\rho}$.

Figure 1.4: Kernel Density Estimates of the Small-Sample Distribution of the Cointegration Test Statistic



Notes: The solid line shows the density of the test statistic in model (1.18) with $\alpha_c = 0.033$. The dotted line describes the density for $\alpha_c = 0.3$ with all other parameters unchanged. Changing α_c moves the 5% quantile from -3.247 to -3.149.

2 Assessing the Anchoring of Inflation Expectations

2.1 Introduction

Expectations play a key role in the conduct of modern monetary policy. In particular, the New-Keynesian Phillips curve stresses the importance of inflation expectations for the rate of actual inflation. Central banks' ability to achieve price stability is thus directly linked to its ability to anchor inflation expectations at their target. Major central banks, including the Federal Reserve, the Bank of England and the European Central Bank, monitor inflation expectations as an indicator of inflation pressure. The quote "inflation expectations are well anchored" is a frequently used phrase in press conferences and monetary policy reports. Yet, in spite of their prominent role in monetary policy, inflation expectations are still under-researched. Specifically, it is not clear how the degree of anchoring of inflation expectations should be defined and measured empirically.

In monetary policy practice it is often argued that inflation expectations are well anchored if their distance to a more or less explicit inflation target is sufficiently small, see BoE (2010) and ECB (2011). More sophisticatedly, the empirical literature employs news regressions and a pass-through criterion. The news regression approach exploits the idea that anchored inflation expectations should be insensitive to economic news, compare Levin et al. (2004) and Gürkaynak et al. (2010b). Similarly, the pass-through criterion of Jochmann et al. (2010) and Gefang et al.

(2012) defines inflation expectations as anchored if longer-term expectations do not respond to changes in shorter-term expectations. Both approaches restrict their attention to first differences of the inflation expectations measure.

In the present paper, we argue that differencing leads to a loss of valuable information and imposes implausible dynamics of inflation expectations. Firstly, a regression in first differences implies a unit root for the level of expected inflation, i.e. shocks to the level never die out. Such extreme persistence appears hardly compatible with the idea of anchored expectations. Secondly, within a first difference regression any information about the level of expected inflation is lost. However, even if the central bank does not announce an explicit inflation target, the level of inflation expectations should be of crucial importance.

We propose an exponential smooth transition autoregressive (ESTAR) model to assess the degree of anchoring. Nobay et al. (2010) recently showed that the ESTAR model captures the dynamics of the actual rate of US inflation remarkably well.[1] As a natural extension, we apply this model to inflation expectations data. The distinguishing feature of the ESTAR approach is given by its flexible dynamics. On the one hand, the model accounts for the locally high persistence typically observed in expectations data, while on the other hand it implies global stationarity, i.e. shocks to the level die out eventually.

The ESTAR model allows inflation expectations to return to some long-run equilibrium value or *anchor*. This value will be interpreted as the *market-perceived* inflation target, which may well deviate from an officially announced inflation target of a central bank. The transition speed within the exponential function determines how fast the reversion to the perceived target takes place and therefore provides a natural measure of the *strength* of the anchor. The transition function of the ESTAR model implies an increasing incentive to revise expectations the more they deviate from the market-perceived target. These characteristics appear suitable and also intuitive in view of anchored inflation expectations generated by credible monetary policy. We include macroeconomic news variables as controls

[1] The ESTAR approach is also used to model the dynamics of other macroeconomic time series as, for instance, exchange rates (Kilian and Taylor 2003).

in the ESTAR model. Therefore, our approach represents an extension of the news regression of Gürkaynak et al. (2010b) which is nested as a special case. The crucial difference is, however, that even if economic news affect inflation expectations, they might still be well anchored. The main aspect of our criterion is, how fast the impact of a shock decays.

We investigate the degree of anchoring of inflation expectations in the United States (US), the European Monetary Union (EMU), the United Kingdom (UK) and Sweden (SW). The expectations measure under consideration is the so called break-even inflation (BEI) rate that is the most prominent measure of inflation expectations within the news regression and the pass-through literature. BEI rates can be derived from the spread of nominal and real government bond yields, i.e. inflation-indexed bonds. Although the considered countries have highly liquid nominal and real bond markets, constant maturity yields of real bonds derived from term structure estimates are usually not readily available. In order to avoid distortions triggered by different data sources and estimation techniques, we closely follow the methodology of Gürkaynak et al. (2010a) and construct a homogeneous data set of BEI rates for the countries under investigation.

With respect to the macroeconomic news variables, we find significant influence in all countries, suggesting equally distorted inflation expectations. The different mean-reverting properties of the expectations series, however, reveal that the degree of anchoring of inflation expectations varies substantially across countries and expectations horizons. We find that shorter-term expectations are anchored more firmly than longer-term expectations, meaning that shocks of a given magnitude die out faster in shorter-term expectations. Among the four countries, the anchoring of expectations is strongest in the EMU, along with a perceived target close to the ECB's implied inflation target of 2%. In contrast, UK inflation expectations exhibit the weakest degree of anchoring, reflecting very high persistence. This is accompanied by a high market-perceived target of up to 4.3%. In view of inflation expectations in the US, a comparison of a pre- and post-Lehman period shows that the strength of the anchor of shorter-term expectations decreases, while it increases for longer-term expectations.

The remainder of the paper is structured as follows. Section 2.2 introduces the anchoring criterion based on the ESTAR model. The measure of inflation expectations, i.e. BEI rates, are introduced in Section 2.3. Furthermore, Section 2.3 comprises a preliminary data analysis. Estimation results, including an impulse response analysis, are discussed in Section 2.4, and Section 2.5 concludes.

2.2 Assessing the Degree of Anchoring

We analyze the degree of anchoring by means of an exponential smooth transition autoregressive (ESTAR) model. Similar ESTAR specifications are prominently used in literature on purchasing power parity, and also to model the actual rate of inflation, see among others Kilian and Taylor (2003) and Nobay et al. (2010), respectively.[2] The model is given by

$$y_t = c + \exp\left(-\gamma(y_{t-1} - c)^2\right)\left(\sum_{i=1}^{p} \alpha_i y_{t-i} - c\right) + \beta X_t + \varepsilon_t \ , \tag{2.1}$$

where y_t represents the measure of inflation expectations, i.e. the BEI rate, and c is a constant. The sum of autoregressive parameters is restricted to $\sum_{i=1}^{p} \alpha_i = 1$ and X_t constitutes a vector of economic news variables, with β as the corresponding coefficient vector and ε_t as a zero mean white noise process. The dynamics of y_t are determined by the exponential smooth transition function $\exp(-\gamma(y_{t-1} - c)^2)$ which is the source of non-linearity in this model. The transition function is bounded between zero and one, and depends on the transition variable y_{t-1}, the smoothness parameter $\gamma > 0$, and a location parameter c. Given the restriction of the sum of autoregressive parameters, model (2.1) behaves locally like a random walk if the lagged expectations measure y_{t-1} equals c. If y_{t-1} departs from c, the process is stationary and the degree of mean reversion at time t depends on the squared difference between y_{t-1} and c. As shown in Kapetanios et al. (2003), the ESTAR model is globally stationary despite its local non-stationarity.

[2] The adequacy of the ESTAR specification in the present context of inflation expectations is evidenced by linearity tests as presented in Appendix 2.B

In economic terms, the equilibrium value c can be interpreted as the *market-perceived* inflation target. If the BEI rate was close to c in the last period, a shock to inflation expectations would have a long lasting impact. That is, the model allows for local high persistence when deviations from the target are so small that they are economically negligible. However, when shocks drive the BEI rate further away from the target, the anchoring pulls expectations back to c. Due to the non-linearity of the model, the persistence of inflation expectations decreases in the distance to the market-perceived target, i.e. the larger the gap between y_{t-1} and c, the stronger the mean reversion. For a given distance, the parameter γ controls the shape of the exponential function and hence the transition speed towards c. Therefore, estimates of γ are considered a natural measure for the *strength of the anchoring* of inflation expectations: the larger the γ, the stronger the anchor. Apart from γ, the market-perceived inflation target c provides further important information, namely the level of expectations, which may or may not be close to the announced central bank's inflation target.

Following the news regression literature, we also include major macroeconomic and monetary news as control variables; compare Gürkaynak et al. (2010b). The crucial difference between our econometric approach and the standard news regression is that model (2.1) allows the expectations measure to mean revert. In fact, the news regression given by

$$\Delta y_t = \beta X_t + \varepsilon_t \tag{2.2}$$

implies a unit root in the level of y_t. The ESTAR extension of (2.2) mitigates this implausible assumption. It provides more flexible inflation expectations dynamics and nests the news regression for the special case $p = 1$ and $\gamma = 0$. Furthermore, the news regression approach in (2.2) is likely to suffer from an omitted variable bias, i.e. due to data availability, the selection of news variables might be incomplete. In contrast, our approach does not rely on the completeness of the X_t vector since the consistently estimated γ reflects the adjustment speed to a shock of any nature, be it a shock in X_t or in ε_t.

2.3 Break-Even Inflation Rates

The expectations data are extracted from nominal and inflation-indexed government bond yields. According to the Fisher equation, the spread between nominal and real yields provides a measure of inflation expectations, i.e. the break-even inflation (BEI) rate. In contrast to a holder of a real bond, the investor of a nominal bond faces inflation risk. Hence, BEI rates are not a pure measure of inflation expectations. In fact, break-even inflation rates consist of inflation expectations and an inflation risk premium. Christensen et al. (2010) show, however, that the average risk premium is virtually zero within an affine term structure model for the US. Furthermore, given our interest in the anchoring of inflation expectations, including the premium as a part of the relevant variable provides important information. More specifically, a central bank that aims to stabilize inflation expectations should also aim to minimize the inflation risk premium. Therefore, we rely on BEI rates to evaluate the anchoring of inflation expectations.[3]

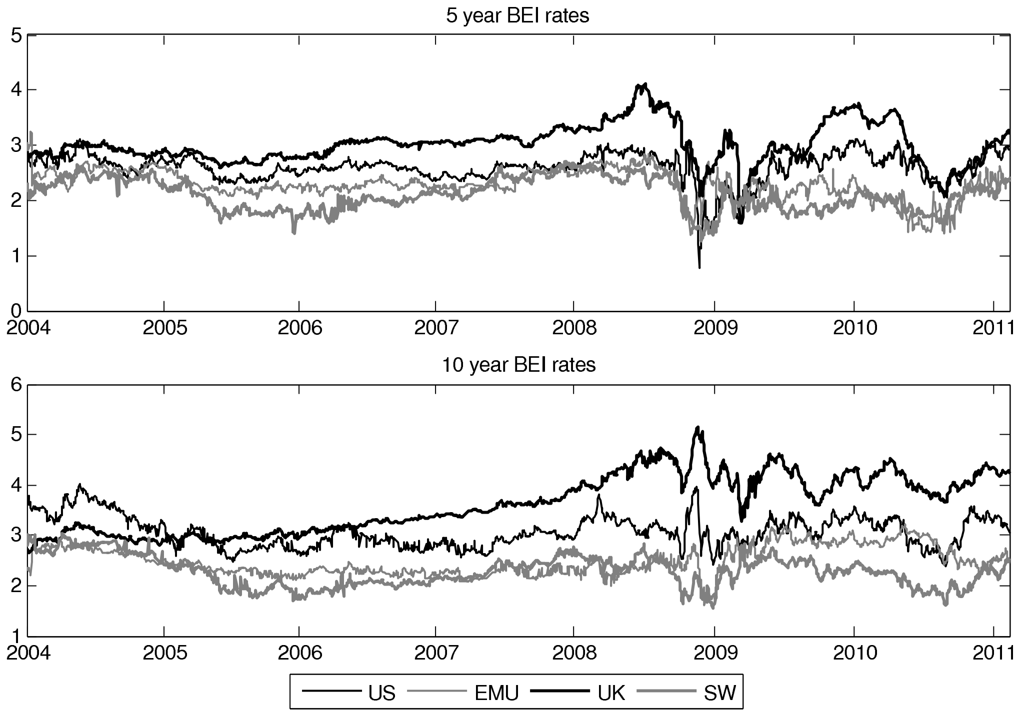
We investigate the degree of anchoring in the United States (US), the European Monetary Union (EMU), the United Kingdom (UK) and Sweden (SW). This selection is consistent with the former anchoring literature and is narrowed by data availability. We investigate daily data from January 2004 to February 2011. By starting in 2004, we ensure that the countries have highly liquid nominal and real bond markets across a wide range of maturities. We follow the methodology of Gürkaynak et al. (2007) and Gürkaynak et al. (2010a) and estimate nominal and real Nelson-Siegel-Svensson yield curves to obtain coherent term structures of BEI rates across the different countries under investigation.[4] The inflation expectations measure is the one-year forward break-even inflation rate. We consider two different expectations horizons: the five-year horizon and the ten-year

[3]In order to disentangle the two components, affine term structure models of e.g. Adrian and Wu (2009) and Christensen et al. (2010) are available. Such a filtered inflation measure, however, strongly depends on the choice of the model and, of course, is subject to estimation uncertainty.

[4]By building a coherent dataset, we intend to minimize the risk of distortions induced by using different data sources that rely on different methodologies. For instance, the FED uses the Nelson-Siegel approach while the Bank of England applies a spline method; the criteria for choosing the specific bonds also differ between different data sources. For details on our methodology see Appendix 2.A.

horizon. The one-year forward in five years captures the ability of central banks to anchor inflation expectations within the often defined policy horizon of three to five years. The ten-year horizon is commonly used in the anchoring literature and represents longer-term expectations, compare Gürkaynak et al. (2010b) and Beechey et al. (2011).

Figure 2.1: Five- and Ten-Year BEI Rates



Notes: Calculated via nominal and real forward rates from Nelson-Siegel-Svensson yield curves. The five-year rate reflects expectations in five years for one year, the ten-year rate expectations in ten years for one year.

Figure 2.1 depicts the five-year (upper graph) and ten-year (lower graph) expectations horizons of the different countries. The figure indicates that market participants expect inflation around two and three percentage points. As expected, revisions of expectations are rather small on a daily basis, such that the depicted series behave very persistently. The degree of anchoring over time is not obvious but an impact of the global financial crisis is clearly visible. Therefore, both a pre-crisis and a crisis period are considered. We follow a standard treatment of

Table 2.1: Descriptive Statistics of Inflation Expectations

		pre-crisis			crisis		
	Horizon	mean	std.	d	mean	std.	d
US	5Y	2.63	0.16	0.75 [0.64,0.86]	2.54	0.40	0.56 [0.40,0.73]
	10Y	3.05	0.32	0.87 [0.76,0.98]	3.11	0.29	0.63 [0.47,0.80]
EMU	5Y	2.40	0.18	0.77 [0.66,0.88]	2.09	0.28	0.52 [0.36,0.67]
	10Y	2.37	0.19	0.63 [0.53,0.74]	2.71	0.31	0.93 [0.76,1.10]
UK	5Y	3.04	0.29	0.90 [0.80,1.02]	2.96	0.46	0.90 [0.74,1.08]
	10Y	3.35	0.48	0.68 [0.57,0.79]	4.15	0.30	1.03 [0.87,1.22]
SW	5Y	2.20	0.28	0.64 [0.54,0.75]	1.93	0.22	0.76 [0.58,0.91]
	10Y	2.33	0.32	0.70 [0.59,0.80]	2.20	0.26	0.82 [0.66,0.99]

Notes: Mean and standard deviation in percentage point and d the order of fractional integration.Pre-crisis sample Jan 2004 - Sep 2008 ($\sim$ 1200 obs.), crisis sample Sep 2008 - Feb 2011 ($\sim$ 630 obs.).

dating the crisis by defining the bankruptcy of Lehman Brothers on 15 September 2008 as its starting point. Note that our crisis period lasts until 2011 and thus incorporates the European sovereign debt crisis as well.

Descriptive statistics of the break-even rates before and after the Lehman bankruptcy are presented in Table 2.1. BEI rates across the different countries and expectations horizons are on average well above the two percent mark. The two percent mark is often viewed as an inflation target in the EMU, the UK and Sweden and very recently in the US as well. The means and standard deviations of most countries are larger for the ten-year horizon than for the five-year horizon. A clear pattern of how the crisis affects the expectations measure is not evident.

To get a first impression of the mean reversion properties of inflation expectations we follow the literature on inflation persistence (e.g., Hassler and Wolters 1995, Meller and Nautz 2012) and estimate the order of fractional integration, d. All estimates lie within the interval $(0.5, 1)$, where processes are still mean reverting but non-stationary.[5] However, it is well-known that estimates of the order of fractional integration are misleading if non-linearities of the true underlying process are not considered, see e.g. Ohanissian et al. (2008) and Kruse and Sibbertsen (2011). In line with this interpretation, we conducted the Kapetanios et al. (2003) and Teräsvirta (1994) linearity tests against ESTAR. These test results provide strong evidence in favor of the ESTAR model (see Appendix 2.B).

2.4 Empirical Results on the Anchoring of Inflation Expectations

2.4.1 Strength and Level of the Anchor

The effect of at least one macroeconomic news variable is statistically significant in all countries under investigation (see Appendix 2.C.2). Strictly applying the news regression criterion, we would conclude, that inflation expectations in the EMU, the US, Sweden and the UK are equally de-anchored. Instead, our extended ESTAR model allows to further analyze shocks to inflation expectations. Specifically, we estimate how long their effect lasts.

The empirical analysis focuses on the two model parameters γ and c in equation (2.1) that provide information on the strength and level of the anchor. Since the ongoing crisis potentially changes the degree of anchoring, we account for parameter shifts within our ESTAR specification (2.1). Specifically, a Lehman step dummy LEH that takes the value one from 9/15/2008 until the end of the sample captures breaks in c and γ.[6] Estimation results on the anchoring of inflation expectations, given in Table 2.2, are interpreted with respect to three perspectives: across the five- and ten-year expectations horizons; across countries; and across the pre-crisis and the crisis period.[7]

[5] The non-stationarity is caused by an unbounded variance of the process.

[6] Note that we found estimation results of the time series dynamics to be robust against a sample split, i.e. separate estimation of the pre-crisis and crisis sample.

[7] Note that the γ and c estimates are robust against the exclusion of the news variables.

Table 2.2: Anchoring of Inflation Expectations

	US		EMU		UK		SW	
	5Y	10Y	5Y	10Y	5Y	10Y	5Y	10Y
c	**2.613**	**3.233**	**2.416**	**2.491**	**3.113**	**3.463**	**2.169**	**2.310**
	(0.029)	(0.059)	(0.021)	(0.035)	(0.108)	(0.140)	(0.049)	(0.066)
LEH	−0.181	**−0.181**	**−0.370**	0.162	0.102	**0.830**	**−0.217**	−0.140
	(0.100)	(0.075)	(0.048)	(0.096)	(0.532)	(0.180)	(0.092)	(0.082)
γ	**0.294**	**0.055**	**0.531**	**0.189**	0.021	0.011	**0.074**	**0.057**
	(0.086)	(0.016)	(0.188)	(0.059)	(0.014)	(0.007)	(0.026)	(0.025)
LEH	**−0.250**	**0.152**	−0.023	−0.113	−0.015	0.020	0.047	0.108
	(0.088)	(0.071)	(0.188)	(0.071)	(0.017)	(0.019)	(0.073)	(0.067)
p	1	4	2	3	2	3	2	2
Q(5)	0.59	0.26	0.95	0.28	0.08	0.05	0.23	0.58
Q(10)	0.60	0.41	0.97	0.35	0.01	0.14	0.18	0.75
ARCH(1)	0.58	0.00	0.86	0.02	0.01	0.10	0.63	0.22
ARCH(5)	0.67	0.11	0.89	0.29	0.07	0.47	0.24	0.56

Notes: $y_t = c + c\,\mathrm{LEH} + \exp\left(-(\gamma + \gamma\,\mathrm{LEH})(y_{t-1} - (c + c\,\mathrm{LEH}))^2\right)\left(\sum_i \alpha_i y_{t-i} - (c + c\,\mathrm{LEH})\right) + \beta X_t + \varepsilon_t$ is estimated by ML. y_t := daily BEI rates are measured in percentage points in the sample period Jan 2004 to Feb 2011. Numbers in bold indicate significance at the 5% level. Bollerslev-Wooldrige heteroskedasticity consistent standard errors of the estimated coefficients are given in parentheses. LEH:= step dummy that takes the value one from 9/15/2008 until 2/14/2011 and zero elsewhere. Estimation results for the news variables can be found in Appendix 2.C.2. The lag length p is determined by standard autocorrelation tests. The p-values of Q-statistics $Q(5)$ and $Q(10)$ illustrate that no significant autocorrelation up to order 10 is left in the residuals. ARCH LM test results reflect that, apart from a few exceptions, the GARCH(1,1) ensures no remaining ARCH effects.

Anchoring Across Expectations Horizons

Estimates of γ are given in the third and fourth row of Table 2.2. The larger the γ, the stronger the anchoring of inflation expectations. Our results indicate a stronger anchoring of five-year BEI rates for all countries under investigation (the US crisis sample excepted). One explanation of this finding might be that the often defined policy horizon of central banks consists of a period between three and five years. This reflects that markets expect a more active role of central banks against medium term inflationary pressure. As a consequence, shocks to longer-term expectations are more persistent. Results on the strength of the anchor are confirmed by estimated market-perceived targets that, on average, take values around 2.5 percentage points for five-year and close to 3 percentage points for ten-year expectations horizon. Deviations of market-perceived targets from a central bank target of two percent can in part be explained by a positive risk premium. In line with Christensen et al. (2010), our findings reflect that markets associate longer-term expectations with higher uncertainty about inflation and thus with a larger premium.

Anchoring Across Countries

For a given level of the market-perceived target, our point estimates of γ suggest that EMU expectations are anchored most firmly, followed by US, Sweden and finally the UK. The transition speed of UK expectations is by far the slowest, reflecting a very high degree of persistence. For the five-year horizon in the UK, for instance, γ equals 0.02. In shorter-term expectations in the EMU, however, we find $\gamma = 0.53$, which indicates a much lower degree of persistence. Considering the location of the anchor, the cross-country comparison reveals that inflation expectations are anchored around the smallest values (between 1.94 and 2.31) in Sweden. For the EMU and US the level of the anchor is located between roughly 2.5 and 3.2 percent. In the UK data on the other hand, we find market-perceived targets of up to 4.29 percent.

Anchoring and the Financial Crisis

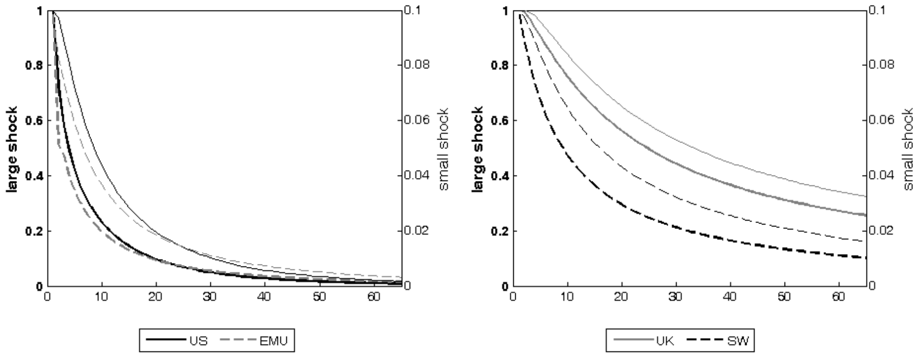
The impact of the crisis does not point in a unique direction. Market-perceived targets, however, change significantly in all countries. In the US and Sweden, c decreases for both expectations horizons, indicating a decrease in inflation pressure. In contrast, the market-perceived target strongly increases by 0.83 percentage points in longer-term UK expectations. In the EMU the term structure of BEI rates steepens due to a decreasing perceived target at the shorter-term horizon (from 2.42 to 2.05) and a non-changing target of longer-term expectations. Unlike the level, the strength of the anchor did not change significantly in most of the countries during the crisis. An exception is given by the US BEI rates. While the transition speed decreases for the medium-term horizon, it increases for the longer-term horizon. This reflects an increasing degree of anchoring of longer-term horizons and a declining degree of shorter horizons.

2.4.2 Impulse Response Analysis

In order to reveal what estimates of γ actually imply for the persistence of inflation expectations, we compute impulse response functions. Since standard techniques for linear processes are not applicable to the ESTAR model, we make use of generalized impulse response functions (GIRFs) as suggested by Koop et al. (1996).[8] The analysis allows us to investigate the anchoring strength for the different countries and time horizons with respect to shocks of different magnitudes. Specifically, non-linear dynamics of inflation expectations are highlighted by two different shock sizes: a *small* shock of 10 basis points (bp) and a *large* shock of 100 bp. Given the estimated standard deviations of the BEI rates in Table 2.1, the small shock roughly represents one half of a standard deviation, while the large shock is approximately four times a standard deviation. In addition to the impulse responses, we calculate half-lives that indicate the number of days an initial shock needs to be absorbed by 50 percent.

[8]See Appendix 2.D for further details on computational steps.

Figure 2.2: Impulse Response Functions of 5-Year BEI Rates



Notes: Generalized impulse response functions are based on ESTAR estimates of the pre crisis as given in Table 2.2. Detailed simulation steps are described in Appendix 2.D. Magnitudes of the shocks are set to 10 bp (*small* shock, indicated by thin lines, right axis) and 100 bp (*large* shock, indicated by bold lines, left axis). The x-axis reflects the number of days after the initial shock.

Figure 2.2 depicts the impulse response functions of the five-year expectations. The figure clearly indicates the nature of the non-linear model: Impulse responses of the large shock decrease much stronger than responses to the small shock. While the initial impact of the large shock drives expectations far away from the perceived target, the anchoring of inflation expectations results in a fast absorption of the shock. In other words, the larger the initial shock, the stronger revisions of expectations in the direction of the perceived target. This effect becomes more pronounced when γ is larger. Half-lives, as reported in Table 2.3, confirm results from the previous subsection. We elaborate along the same three perspectives: across expectations horizons, across countries and through pre-crisis and crisis sample.

Half-lives indicate the better anchoring of shorter-term expectations. In comparison to the five-year horizon, a shock to inflation expectations at the ten-year horizon needs, on average, about one week more to be absorbed by 50%.

Turning to the cross-country comparison, half-lives of up to 4 weeks mirror the strong anchoring of EMU, Sweden and US expectations. In contrast, the time to absorb 50% of a shock is 5 to 10 weeks in the UK. The overall impression from

Table 2.3: Half-lives in Days

		pre-crisis		crisis	
	Horizon	10 bp	100 bp	10 bp	100 bp
US	5Y	9	4	19	12
	10Y	19	16	9	5
EMU	5Y	7	3	6	2
	10Y	11	7	9	5
UK	5Y	34	25	28	24
	10Y	43	40	49	42
SW	5Y	16	10	17	11
	10Y	16	11	18	10

Notes: Reported values represent the absorption time measured in days for 50% of the initial shock size of 10 bp ("small shock") and 100 bp ("large shock"), respectively.

Table 2.3 is, however, that shocks are absorbed fairly rapidly in all four countries, which expresses a substantial degree of anchoring.

In view of the pre-crisis and crisis period, half-lives of US inflation expectations reflect the significant break in the anchoring strength. The break reverses the absorption time of longer- and shorter-term expectations in that shorter horizons appear less anchored during the crisis. Since the expectations data for the EMU, the UK and Sweden display no significant break in γ, the slight changes in half-lives simply result from the particular set of histories used to compute impulse responses in each subsample.

In sum, the decay of the impulse responses illustrates the main idea of the proposed anchoring measure: well-anchored inflation expectations should display stationary characteristics and should therefore always tend to return to some long-run equilibrium value.

2.5 Conclusion

In this paper we propose a non-linear time series approach to determine the degree of anchoring. We build on a well-known analytical framework for investigating this topic, namely the news regression, wherein first differences of an expectations measure are regressed on macroeconomic news. Specifically, the implicit unit root assumption of the news regression approach is relaxed as we allow anchored inflation expectations to follow a globally stationary ESTAR process. This generalization permits a shift in focus from the immediate news effect to the dynamics of inflation expectations. Model parameters are economically interpretable as a market-perceived inflation target and as the strength of the anchor that drives expectations back to the target.

Macroeconomic news variables turn out to have a significant impact on inflation expectations in the US, EMU, UK and Sweden. While in the news regression context this result suggests equally distorted and non-stationary expectations, the proposed ESTAR extension reveals mean-reversion and thus well anchored inflation expectations in all countries under investigation. The ESTAR estimates show, firstly, that shorter-term inflation expectations are anchored more firmly than longer-term expectations. Secondly, expectations appear best anchored in the EMU, followed by the US, Sweden and the UK. Thirdly, in most countries the average level of inflation expectations decreases during the crisis. Given central banks' mandate to stabilize the actual rate of inflation, our results support the view of credible policy strategies that anchor inflation expectations in all countries investigated here. Apart from the UK, the market perceived inflation targets are close to the usually imposed inflation targets of around 2 percent. Moreover, expectations vary around these targets in a stationary manner. This leads to the conclusion that expectation formation processes are successfully controlled by central banks.

So far, the univariate setup implies that shocks to inflation expectations are uncorrelated across countries. Of course, for many cases, such as oil price shocks, the uncorrelatedness assumption might be too rigid. A multivariate extension of the present approach would therefore provide an interesting path for future research.

2.A Term Structure Estimation

This appendix describes how the constant maturity series were estimated as well as which and how many bonds were used. Our data set is based on bond yields downloaded from the Bloomberg database (for an overview see Table 2.4). We take all outstanding, standard government bonds (e.g. no callables) with more than two years time to maturity into account. Bonds with less than 24 months to maturity are cut out because real bonds' indexation lags erraticate the yields of these securities, as shown in Gürkaynak et al. (2010a).

Table 2.4: Nominal and Real Bonds - Overview

Short name	indexation	$\#'04$	$\#'10$	Obs.
US TREASURY N/B	-	124	204	1792
TSY INFL IX	US CPI	12	31	
FRANCE O.A.T.	-	31	34	1858
FRANCE O.A.T.I/L	EMU HICP	4	6	
UK TREASURY	-	26	35	1856
UK TSY I/L	UK CPI	9	17	
SWEDEN GOVT	-	12	9	1837
SWEDEN I/L	CPI Sweden	5	5	

Notes: $\#'04$, $\#'10$ report the number of outstanding bonds in June 2004 and 2010, Obs. is the number of daily observations within the time span of January 2004 until February 2011. EMU HICP refers to the harmonized index of consumer prices of the European Monetary Union.

For each day where yields for more than three bonds are available, we follow the approach of Gürkaynk et al. (2007), and estimate constant maturity yields. The standard parametric yield curve specification is based on a functional form that was proposed by Nelson and Siegel (1987) and extended by Svensson (1994):

$$\hat{z}_t(\tau) = \beta_1 + \beta_2 \left[\frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}} \right] + \beta_3 \left[\frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}} - e^{-\frac{\tau}{\lambda_1}} \right] + \beta_4 \left[\frac{1 - e^{-\frac{\tau}{\lambda_2}}}{\frac{\tau}{\lambda_2}} - e^{-\frac{\tau}{\lambda_2}} \right] .$$

The observed zero coupon yield for maturity τ is given by $z(\tau)$, whereas the model-implied yield is $\hat{z}(\tau)$. We minimize $\sum(\hat{z}(\tau) - z(\tau))^2$ with respect to the parameters β_1, β_2, β_3, β_4, λ_1 and λ_2 by the Differential Evolution approach proposed by Storn and Price (1997). Forward rates are derived from zero-coupon yield curves via

$$f_t(n,m) = \frac{1}{m}((n+m)\hat{z}_t(n+m) - n\hat{z}_t(n)) \, , \tag{2.3}$$

in which $f_t(n,m)$ is the forward rate at time t for a period of m years, beginning n years in the future. The n-year BEI rate reflects today's expected inflation rate (plus an inflation risk premium) and is given by $\text{BEI}(n) = f_t^{\text{nom}}(n,m) - f_t^{\text{real}}(n,m)$.

2.B ESTAR Specification Tests

For specification of the ESTAR model, we perform two different types of linearity tests. Both approximate the exponential function in (2.1) around $\gamma = 0$ to obtain an auxiliary regression. The t-test of Kapetanios et al. (2003) (KSS) tests the null of a linear non-stationary autoregressive process against ESTAR non-linearities. We also run the LM test of Saikkonen and Luukkonen (1988) and Teräsvirta (1994) to test the null of a linear, stationary autoregressive process against ESTAR non-linearities. Both tests are carried out since conventional autoregressive models are close to non-stationarity up to such a degree, that standard unit root tests show conflicting test results.

Table 2.5 shows the results of the two linearity tests. The LM test rejects the null across all countries and almost all sample periods. For a given country, an expectation horizon and a sample period, at least one of the two tests rejects the null in favor of the ESTAR model.[9] Consequently, we interpret these results as

[9] The exception is the Swedish five-year BEI rate in the crisis sample. However, the test statistic of 7.41 corresponds to a p-value of 0.11 and the KSS test statistic of -2.33 falls above the 10% level of -2.66. Even though tests fail to reject the null at the 10% level, they point in the non-linear direction. In general, due to the sample split, linearity tests tend to suffer from low power, which may partly explain the failure of rejecting the null in a few cases.

Table 2.5: Linearity Tests Against ESTAR

		H_0 : stationary AR(p)			H_0 : non-stationary AR(p)		
	Horizon	whole	pre-crisis	crisis	whole	pre-crisis	crisis
US	5Y	41.4^{***}	21.3^{***}	10.6^{***}	-5.19^{***}	-4.10^{***}	-3.16^{**}
	10Y	33.4^{***}	44.5^{***}	7.08	-4.01^{***}	-2.17	-2.91^{**}
EMU	5Y	59.5^{***}	65.2^{***}	6.70	-5.52^{***}	-10.8^{***}	-3.15^{**}
	10Y	44.9^{***}	44.9^{***}	16.9^{**}	-2.14	-2.11	-1.96
UK	5Y	33.4^{***}	44.5^{***}	5.34	-3.99^{***}	-1.32	-2.75^{*}
	10Y	46.6^{***}	44.2^{***}	42.9^{***}	-2.61	-1.37	-1.72
SW	5Y	11.9^{**}	21.9^{***}	7.41	-3.59^{***}	-3.11^{**}	-2.33
	10Y	38.4^{***}	30.6^{***}	33.4^{***}	-4.55^{***}	-3.43^{***}	-4.22^{***}

Notes: Test statistics of the LM test with the null hypothesis of a stationary autoregressive process and KSS with the null of a non-stationary autoregressive process. The lag length is chosen in that residuals are free from significant autocorrelation. The rejection of the respective null hypothesis at the 10% is indicated by *, at the 5% by ** and at the 1% level by ***. Sample periods: whole refers to January 2004 - February 2011; pre-crisis to January 2004 - September 2008; and crisis to September 2008 - February 2011.

conclusive evidence that the true underlying processes can be well-approximated by an ESTAR model.

2.C News Variables

2.C.1 News Data

The news variables are calculated by the difference between the actual and the expected value. The expected value is represented by the mean prediction of the Bloomberg survey of professional economists, mostly consisting of bankers. They submit their forecast before or on the Fridays prior to the data release. The actual and forecasted values of the advanced estimate of the gross domestic product (GDP), industrial production (IP), consumer price index (CPI) and the producer price index (PPI) refer to the percentage yearly basis. The GDP, IP, CPI and PPI news measure the difference between the actual and forecasted value in percentage points. The unemployment rate (UMP) and the monetary policy rate (MP) are measured in percentages. The respective news variable reflects the unexpected component in percentage points. In line with the rational expectations assumption, mean forecast errors are close to zero, mostly uncorrelated and some do not reject the null of normality.[10]

2.C.2 Estimation Results for News Variables

While time series dynamics determine the degree of anchoring, surprise components of major economic announcements reveal potential sources of shocks that drive expectations away from the market-perceived target. Estimation results on the news coefficients are reported in Table 2.6. Numbers in bold reflect significance at the 10% level, indicating news announcements that lead to systematic revisions in inflation expectations. Monetary policy news, for example, show a significant impact in all countries. In general, for each country we observe at least one announcement that move markets' expectations significantly.

[10] We run the same type of regressions with median expectations. Qualitative results remain the same.

Table 2.6: Estimation Results for News Variables

	US		EMU		UK		SW	
	5Y	10Y	5Y	10Y	5Y	10Y	5Y	10Y
GDP	0.006	0.002	**0.088**	−0.013	0.025	−0.002	−0.004	−0.003
	(0.006)	(0.013)	(0.034)	(0.032)	(0.028)	(0.015)	(0.010)	(0.006)
IP	**0.023**	−0.019	0.002	−0.004	−0.002	−0.003	0.002	0.001
	(0.010)	(0.015)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)
UEM	**−0.070**	0.049	**−0.074**	−0.054	0.037	−0.007	0.001	0.051
	(0.037)	(0.031)	(0.045)	(0.047)	(0.030)	(0.048)	(0.016)	(0.043)
CPI	−0.023	−0.029	0.058	−0.110	0.085	**0.043**	**−0.050**	0.018
	(0.028)	(0.037)	(0.079)	(0.110)	(0.055)	(0.026)	(0.023)	(0.024)
PPI	0.006	0.001	0.023	0.028	−0.011	0.010	0.011	0.007
	(0.004)	(0.013)	(0.618)	(0.024)	(0.099)	(0.008)	(0.009)	(0.010)
MP	**−0.280**	**−0.469**	0.182	**0.657**	−0.021	**−0.048**	0.044	**0.167**
	(0.063)	(0.070)	(0.414)	(0.349)	(0.033)	(0.018)	(0.059)	(0.084)

Notes: ML estimation results for the news variables X_t in equation (2.1), $y_t = c + c\,\text{LEH} + \exp\left(-(\gamma + \gamma\,\text{LEH})(y_{t-1} - (c + c\,\text{LEH}))^2\right)\left(\sum_i \alpha_i y_{t-i} - (c + c\,\text{LEH})\right) + \beta X_t + \varepsilon_t$. Daily BEI rates, y_t, are measured in percentage points in the sample period January 2004 to February 2011. Bollerslev-Wooldrige heteroskedasticity consistent standard errors of the estimated coefficients are given in parentheses. Symbols in bold indicate significance at the 10% level. c and γ estimates are reported in Table 2.2.

2.D Generalized Impulse Response and X-Life

2.D.1 Impulse Response

In order to calculate the GIRFs, we follow Koop et al. (1996). The GIRF at $t+h$ is defined as the difference between the expected value of a stochastic process conditioned on an impulse ξ hitting the process at time t, and the conditional expectation that is obtained without such a shock:

$$\mathrm{GIRF}(h,\xi,\omega_{t-1}) = \mathbb{E}[y_{t+h}|y_t+\xi,\omega_{t-1}] - \mathbb{E}[y_{t+h}|y_t,\omega_{t-1}] \ , \tag{2.4}$$

where ω_{t-1} refers to one particular history of the process y_t. $\mathrm{GIRF}(h,\xi,\omega_{t-1})$ represents one realization of the random variable $\mathrm{GIRF}(h,\xi,\Omega_{t-1})$ and can be approximated via stochastic simulation. To calculate $\mathbb{E}[y_{t+h}|y_t+\xi,\omega_{t-1}]$ and $\mathbb{E}[y_{t+h}|y_t,\omega_{t-1}]$, we average over 1000 future paths, in which each y_{t+h} is created by iterating it on the ESTAR model with parameter values equal to those from the empirical estimates and with randomly drawn GARCH(1,1) errors with i.i.d. normal innovations. The impulse ξ is set to the size of one residual standard deviation, i.e. $\xi=\sigma_\varepsilon$. The aspect of interest of the random variable $\mathrm{GIRF}(h,\sigma_\varepsilon,\Omega_{t-1})$ is given by its unconditional mean:

$$\mathbb{E}[\mathrm{GIRF}(h,\sigma_\varepsilon,\Omega_{t-1})] = \mathbb{E}[y_{t+h}|y_t+\sigma_\varepsilon] - \mathbb{E}[y_{t+h}|y_t] \ . \tag{2.5}$$

We approximated equation (2.5) by averaging over all ω_{t-1} observed in the sample. Note that it is the unconditional mean of the GIRF that we refer to simply as GIRF or *impulse response* throughout the paper.

2.D.2 X-Life

Following Dijk et al. (2007), X-lives are estimated by:

$$X\text{-life}(x,\sigma_\varepsilon) = \sum_{m=0}^{\infty}\left(1-\prod_{h=m}^{\infty} \quad (x,h,\sigma_\varepsilon)\right) \ , \text{with} \tag{2.6}$$

$$\begin{aligned}(x,h,\sigma_\varepsilon) = \quad & \Big(\Big|\mathbb{E}[\mathrm{GIRF}(h,\sigma_\varepsilon,\Omega_{t-1})] - \lim_{h\to\infty}\mathbb{E}[\mathrm{GIRF}(h,\sigma_\varepsilon,\Omega_{t-1})]\Big| \\ & \leq x\Big|\sigma_\varepsilon - \lim_{h\to\infty}\mathbb{E}[\mathrm{GIRF}(h,\sigma_\varepsilon,\Omega_{t-1})]\Big|\Big) \ .\end{aligned}$$

$0\leq x\leq 1$ refers to the chosen fraction of noise absorption ($x=0.5$ and $x=0.75$ in the application) and $\quad(\cdot)$ is the indicator function.

3 Testing the Preferred-Habitat Theory: The Role of Time-Varying Risk Aversion

3.1 Introduction

A key question for monetary policy is how to effectively influence longer-term yields in order to control inflation or to provide stimulus to aggregate demand. One possible answer is to alter the maturity structure of government debt. This view is advocated by the preferred-habitat approach, which received increasing attention in a series of recent papers (Vayanos and Vila 2009, Greenwood and Vayanos 2010, Greenwood and Vayanos 2012 and Guibaud et al. 2013). The basic idea of the preferred-habitat theory is that investor clienteles with preferences for certain maturities play a crucial role in the determination of bond yields. The main theoretical implication is a positive relation between yields and the relative supply of longer-term debt. The literature emphasizes, however, an important qualification of this prediction. The strength of the positive relation depends on the risk aversion of arbitrageurs that participate in the bond market and undo the preferred-habitat effects.

Despite the growing theoretical literature, empirical evidence on a relation between debt and bond yields is limited. Reinhart and Sack (2000) find the term spread to be negatively related to the government surplus, indicating that debt supply affects the yield curve. Bernanke, Reinhart and Sack (2004) and Greenwood and Vayanos (2010) provide some descriptive results on how bond yield movements in the US may be attributed to changes in the maturity structure of government debt. On days where long-term debt purchases or ceasing of new issuance are announced, figures of yield spreads display a distinct decline.

This paper builds on the approach of Greenwood and Vayanos (2012) who show in a regression analysis of US data that the impact of relative longer-term debt supply on term spreads is economically and statistically significant. In static regressions with constant coefficients, spreads are found to react by up to 38 basis points to a one standard deviation increase in longer-term debt. Even though the standard regression framework constitutes the natural empirical starting point, it might be too restrictive to test for preferred-habitat effects. Therefore, we propose to extend the approach in two dimensions.

Firstly, the preferred-habitat theory implies that the impact of debt supply on yield spreads is stronger when risk aversion of arbitrageurs is high. Imposing constant coefficients rules out any state-dependency of the relation a priori. Secondly, bond yields and relative supply of long-term debt are typically very persistent. A static model does not control for serial correlation and may therefore produce spurious results.

We show that it is essential to take both aspects into account when the preferred-habitat theory is analyzed empirically. On the one hand, we propose an augmented distributed lag (ADL) model. The ADL model avoids the risk of spurious results, even in case of extremely persistent time series. On the other hand, we allow the effect of debt supply on spreads to depend on the state of risk aversion. Thereby, risk aversion is proxied by bond market volatility.

Over a variety of asset pricing models, there is agreement on counter-cyclical risk aversion which increases when marginal utility is high and decreases when marginal utility is low, see Campbell and Cochrane (1999), Rosenberg and Engle (2002) or Gordon and St-Amour (2004). Market volatility also features counter-cyclical movements. It is commonly known to be higher in bad than in good times which makes a connection to risk aversion intuitive. In fact, it is often elaborated on a theoretical relation between risk aversion and market volatility. Mele (2007) and Aydemir (2008), for instance, argue that counter-cyclical risk aversion is the major driving force of volatility, since rational asset evaluation depends on the current state of the economy. Above all, our risk aversion proxy fits the conjecture of Gürkaynak and Wright (2012) that one should observe more pronounced preferred-habitat effects in turbulent than in normal times.

The present paper applies these considerations to the empirical analysis of the preferred-habitat theory and considers volatility as a natural proxy of risk aversion. To this end, we use the most simple candidate of market volatility that can be extracted directly from yield data, i.e. the GARCH variance of the term spread at time t. Methodologically, this amounts to a dynamic regression with the conditional variance entering the mean equation to govern the state-dependency of the effect of longer-term debt supply on spreads.

The analysis is based on daily observations of German government bonds. While constant maturity series of yields can be easily obtained, data on the maturity structure of debt are not readily available. Therefore, this paper generates a new data set of relative debt supply constructed from daily bond prices. At any point in time, the data contain all future coupon and principal payments due within a certain period.

Our empirical results indicate the following. First, estimates from a static regression indicate a significant constant impact of relative supply of longer-term debt on yield spreads. The estimated coefficients are remarkably similar in magnitude to those obtained in previous studies of monthly US data. In a dynamic regression, however, these effects turn out to be spurious. Second, the introduction of state-dependent coefficients reveals strong evidence that a relation between spreads and debt actually exists. Most importantly, this relation survives in the ADL specification. The reaction of spreads to a one standard deviation increase in longer-term debt supply ranges from 5 basis points in times of low risk aversion to 33 basis points when risk aversion is high.

The rest of the paper proceeds as follows. The next section briefly reviews the preferred habitat-model and states the testable hypotheses. Section 3.3 introduces the econometric methodology. The data on German bond yields and relative supply of longer-term debt are presented in Section 3.4. Section 3.5 discusses the empirical results and section 3.6 concludes.

3.2 The Preferred-Habitat Theory

The key implication of the preferred-habitat model of Greenwood and Vayanos (2012) is that the term spread should react positively to changes in the relative supply of longer-term debt. The reaction of the spread, however, is supposed to be stronger when risk aversion is high. To see this, we initially review the main aspects of the model and then turn to the intuition behind the theoretical predictions.

3.2.1 The Model Greenwood and Vayanos (2012)

The yield of a τ-year bond is determined by the interaction between three types of agents: the government, investors with a preference for maturity τ[1] and arbitrageurs. The gross supply of τ-year bonds through the government less the demand of preferred-habitat investors results in a net supply, $\mathrm{NS}_t^{(\tau)}$, at that specific maturity. The time t value of net supply is assumed to be negatively related to the yield $y_t^{(\tau)}$:

$$\mathrm{NS}_t^{(\tau)} = \psi(\tau) - \omega(\tau)\tau y_t^{(\tau)} \quad . \tag{3.1}$$

The constant $\psi(\tau)$ and the slope parameter $\omega(\tau)$ are some functions of τ, with the only restriction that $\omega(\tau) > 0$. The negative dependency on the yield is motivated as follows. First, a higher yield would raise the demand of preferred-habitat investors. Second, if yields increase, prices decrease. Both effects have a negative impact on the value of net supply.

[1]Investors with a preference for shorter maturities are typically associated with banks who prefer to stay liquid whereas demand at longer maturities is often ascribed to insurance companies or pension funds.

For the market to clear, $\mathrm{NS}_t^{(\tau)}$ has to be absorbed by the demand of arbitrageurs, $x_t^{(\tau)}$. They aim for high mean and low variance of their wealth changes $\mathrm{d}W_t$:

$$\max_{\{x_t^{(\tau)}\}_{\tau\in(0,T]}} \left[\mathbb{E}_t(\mathrm{d}W_t) - \frac{a}{2}\mathrm{Var}_t(\mathrm{d}W_t)\right] \quad . \tag{3.2}$$

The remaining part of the model follows the standard Vasicek (1977) setup: The short-rate is the only source of uncertainty in the model and its dynamics are Ornstein-Uhlenbeck. Bond prices are assumed to be affine functions of the short rate.

In equilibrium, it can be shown that the risk premium $\theta_t^{(\tau)}(a)$ for holding a τ-year bond is given by the product of the bond's sensitivity to short rate risk, $A(\tau, a)$, and the market price of risk $\lambda(a)$:

$$\theta_t^{(\tau)}(a) = A(\tau, a)\,\lambda(a) \quad . \tag{3.3}$$

The parameter a in (3.2) and (3.3) refers to the degree of arbitrageurs' risk aversion and constitutes a decisive element to qualify the predictions of the preferred-habitat theory. To see this, it is important to note that any preferred-habitat effect, i.e. any response of yields to changes in bond supply, occurs through the risk premium. Without risk aversion, there are no preferred-habitat effects.

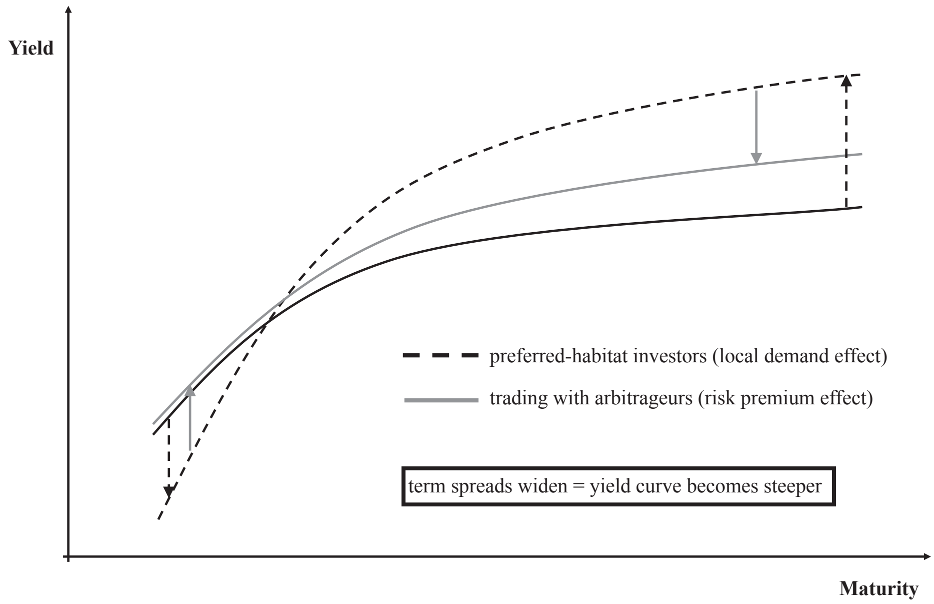
3.2.2 Testable Hypotheses

The testable hypotheses can be derived from the equilibrium term structure of the model. All formal proofs are given in Greenwood and Vayanos (2012). In the following, assume that risk aversion a is positive and constant.

Hypothesis 1: Changes in Debt Supply
The term spread between the yield of τ-year bond and the short-rate is increasing in the relative supply of longer-term debt. The effect is stronger for larger τ.

To see the intuition behind this prediction, suppose that the relative supply of longer-term debt increases. According to Greenwood and Vayanos (2012), this is modeled as a decrease in the constant term $\psi(\tau)$ of the net supply equation (3.1) for small τ and an increase for large τ. The consequences for equilibrium yields are best described if the bond price formation process is thought of as being sequential.

Figure 3.1: Reaction of Yields to Supply Shocks



Notes: The solid black line represents the yield curve at some arbitrary day. In absence of arbitrageurs, as a response to a shock to relative supply of longer-term debt, local preferred-habitat demand causes shorter-term yields to decrease and longer-term yields to increase, i.e. the yield curve rotates. The new yield curve is given by the dashed black line. Arbitrageurs react to the rotation by buying long-term bonds and selling short-term bonds. Thereby, the risk premium increases. Because of the higher premium, trading across maturities raises shorter-term yields even above the solid black line and pushes longer-term yields below the dashed black line. The new yield curve, given by the solid gray line, is the result of an upward shift and a counter-clockwise rotation.

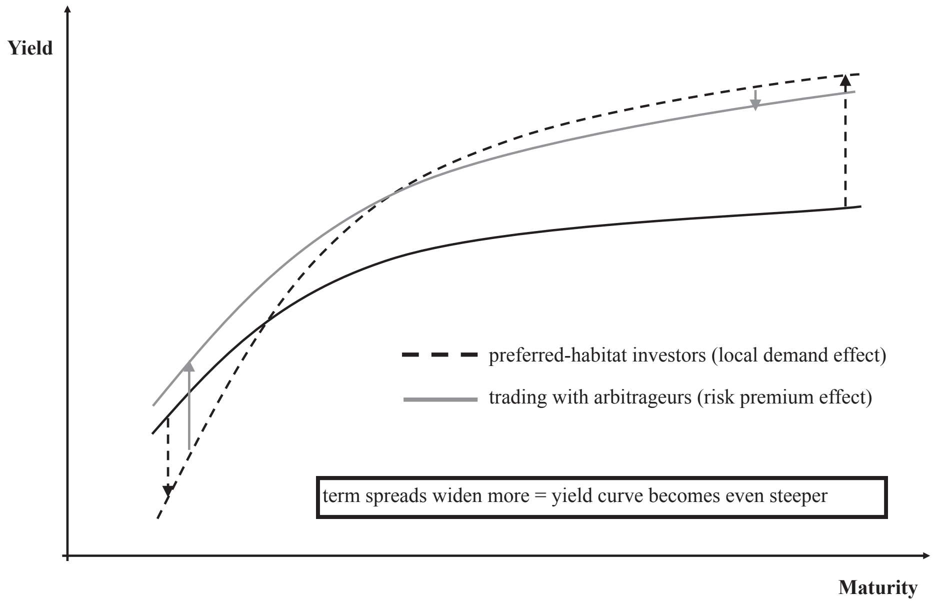
If there were no arbitrageurs, yields would be solely determined by (3.1) and the market of τ-year bonds would clear for $y_t^{(\tau)} = \psi(\tau)/\omega(\tau)\tau$. Therefore, longer-term bond yields increase while shorter-term yields decrease. In Figure 3.1 this is illustrated by a rotation of the yield curve from the solid black line to the dashed black line. Arbitrageurs would now exploit the differences in yields by selling longer-term and buying shorter-term bonds. They thereby tend to reverse the initial changes in yields. As a matter of fact, however, the risk exposure of arbitrageurs has increased since the reshuffling of their portfolios implies that they hold a larger amount of longer-term bonds. This leads to a higher market price of risk and thus to an increase in risk premia at all maturities. The increase in risk premia, in turn, reduces prices and raises yields. Since the sensitivity of bonds to short-rate risk is higher for longer maturities, the increase in premia is stronger for longer-term bonds. Thus, the rise in yields is more pronounced for longer maturities and term spreads between τ-year bonds and short-term bonds widen. The solid gray line in Figure 3.1 represents the new equilibrium yield curve after a shock to the relative supply of longer-term debt.

Hypothesis 2: Changes in Risk Aversion
When arbitrageurs are more risk averse, the effect of longer-term debt supply on spreads is stronger for all τ.

Suppose that the risk aversion of arbitrageurs changes. It is a central comparative static result of the model that the response of the term spread to an increase in the relative supply of debt is stronger when a is high. The idea behind this result can be explained as follows. In the extreme case where arbitrageurs are risk neutral, i.e. $a = 0$, the market price of risk would be zero. Local effects of supply changes would be completely offset by arbitrageurs so that yields would remain unchanged. In fact, bond yields are fully determined by expectations of arbitrageurs on future short-rate developments. In the other extreme case, where arbitrageurs are infinitely risk averse, i.e. $a \to \infty$, risk premia would go to infinity and arbitrageurs would not participate in the market at all. Instead, bond markets would be completely segmented and yields would be fully determined by local demand and supply. For all intermediate cases, the rise in risk premia caused by an increase in the average maturity of arbitrageurs' portfolios, is stronger when risk aversion is high. This is because both components of the premium in (3.3), that is, the sensitivity of bonds to risk and the market price of risk, are increasing in a. A comparison of Figure 3.1 to a situation of higher risk aversion in Figure 3.2

illustrates this mechanism. The steepening of the yield curve is more pronounced when risk aversion is high.

Figure 3.2: Reaction of Yields to Supply Shocks Under *High Risk Aversion*



Notes: This figure shows a state where *risk aversion* of arbitrageurs is *high.* The solid black line represents the yield curve at some arbitrary day. In absence of arbitrageurs, as a response to a shock to relative supply of longer-term debt, local preferred-habitat demand causes shorter-term yields to decrease and longer-term yields to increase, i.e. the yield curve rotates. The new yield curve is given by the dashed black line. Arbitrageurs react to the rotation by buying long-term bonds and selling short-term bonds. Thereby, due to the *high* risk aversion, the risk premium increases *considerably.* Because of the higher premium, trading across maturities raises shorter-term yields *well above* the solid black line and pushes longer-term yields only *slightly below* the dashed black line. The new yield curve, given by the solid gray line, is the result of an upward shift and a counter-clockwise rotation.

3.3 Econometric Methodology

3.3.1 The Static Regression

To estimate the effect of relative supply of longer-term debt on yield spreads, Greenwood and Vayanos (2012) propose the following regression:

$$s_t^{(\tau)} = \beta_0 + \beta_1 D_t + u_t \quad . \tag{3.4}$$

Here, $s_t^{(\tau)}$ denotes the spread between a τ-year bond and the short-rate and D_t refers to the value of longer-term debt supply relative to the total value of debt.[2]

The regression in (3.4) is considered the natural starting point. Since the theoretical model also assumes exogeneity of debt supply, we adopt this assumption throughout the empirical analysis. However, the approach in (3.4) is extended in two respects. Firstly, in order to ensure sound inference, we propose a dynamic regression. Secondly, to capture changes in risk aversion, we allow the response of $s_t^{(\tau)}$ to D_t to be state-dependent and do not impose the restriction that β_1 is constant.

3.3.2 Introducing Dynamics

The dynamic version of (3.4), including lagged values of both variables, only constitutes the straightforward transformation if autocorrelation is present in u_t. In that case, inference in a static regression can be severely biased and produce spurious results. Note that it is very likely for the u_t's to be serially correlated since yields spreads and relative debt supply represent two highly persistent time series.[3] The corresponding extension of (3.4) is given by

$$s_t^{(\tau)} = \beta_0 + \beta_1 D_t + \psi(L) D_{t-1} + \phi(L) s_{t-1}^{(\tau)} + \epsilon_t \quad . \tag{3.5}$$

In (3.5), $\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + ... + \psi_{r-1} L^{r-1}$ and $\phi(L) = \phi_0 + \phi_1 L + \phi_2 L^2 + ... + \phi_{p-1} L^{p-1}$ represent polynomials in the lag operator L. In practice, lag orders p

[2]The measuring of D_t is discussed in detail in the next section.

[3]Appendix 3.B discusses the borderline case of extreme persistence.

and r are chosen so that residuals are white noise and standard inference can be applied. Note that the overall impact of D_t on $s_t^{(\tau)}$ in the ADL model is given by $[\beta_1 + \psi(1)] \cdot [1 - \phi(1)]^{-1}$.

3.3.3 How To Proxy Risk Aversion

The specification in (3.5) rules out any state-dependency of the impact of relative longer-term debt supply on the spread. Therefore, we drop the restriction that β_1 is constant and allow the coefficient to depend on risk aversion.

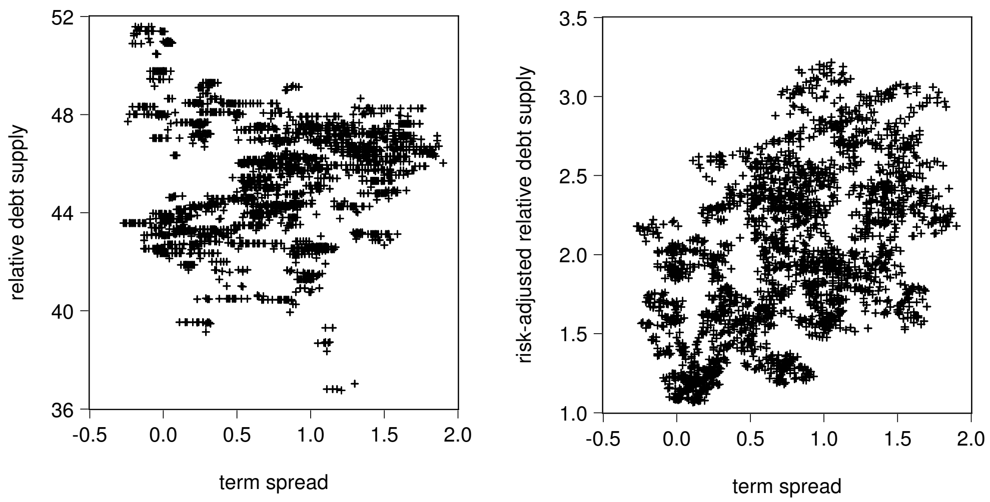
According to the literature on time-varying risk preferences (e.g. Campbell and Cochrane 1999, Rosenberg and Engle 2002 or Gordon and St-Amour 2004), risk aversion is supposed to be high (low) in precisely those periods when marginal utility is also high (low). This counter-cyclical property of risk aversion is also reflected in market volatility. In bad times, we usually observe higher volatility than in good times. Mele (2007) and Aydemir (2008) support this connection, stating that time-varying risk aversion is the major driving force behind counter-cyclical volatility.

We follow this literature and consider volatility as a reasonable choice to proxy risk aversion. This leaves us with the question of a meaningful measure of bond market volatility. To approach this problem, we examine the variability of shocks to the slope of the yield curve, i.e. to the spread $s_t^{(\tau)}$. The slope represents a central summary statistic of the bond market and the shocks to it can be approximated directly from the data that is being researched here by taking the first differences of $s_t^{(\tau)}$.

In order to provide a rough idea of the current state of risk aversion, we consider the rolling standard deviation of $\Delta s_t^{(\tau)}$ over some period, say $\hat{\sigma}_t^{\text{roll}}$ over one quarter. If a state-dependent relation between $s_t^{(\tau)}$ and D_t actually exists, it is natural to analyze a linear dependency as a first approximation. Figure 3.3 shows two scatter plots. The first one plots $s_t^{(5)}$ against D_t, whereas the second one plots $s_t^{(5)}$ against

$\hat{\sigma}_t^{\text{roll}} \cdot D_t$, the relative supply of longer-term debt adjusted by risk aversion.[4] It is difficult to determine whether there is any relation by only looking at the first plot. The observation pairs in the second plot, however, clearly indicate that time-varying risk aversion may reveal a significantly positive relation and thus represents a decisive element in the analysis of the preferred-habitat theory.

Figure 3.3: The Term Spread Against Unadjusted and Adjusted Relative Supply of Longer-Term Bonds



Notes: The first picture shows a plot of $s_t^{(5)}$ against D_t whereas a plot of $s_t^{(5)}$ against $\hat{\sigma}_t^{\text{roll}} \cdot D_t$ is shown in the second picture. $\hat{\sigma}_t^{\text{roll}}$ refers to the rolling standard deviation of changes in the term spread with a window of one quarter. Relative supply of longer-term debt is denoted by D_t, measuring the value of debt that has to be paid in 5 years hence or later relative to the total value of debt.

Because of the intention to conduct a thorough, empirical investigation, we apply a more sophisticated measure than the rolling standard deviation, i.e. the GARCH variance. A GARCH still represents a fairly simple candidate to estimate time t volatility and can be integrated into a tractable time-varying coef-

[4]We choose $s_t^{(5)}$ arbitrarily as a representative example. Since we later analyze several different yield spreads, it is noted that any of them generates almost the same scatter plots as in Figure 3.3.

ficient framework. Building on the dynamic regression in (3.5), we propose the following ADL-GARCH-M model[5]:

$$s_t^{(\tau)} = \beta_0 + \beta_{1t} D_t + \psi(L) D_{t-1} + \phi(L) s_{t-1}^{(\tau)} + \epsilon_t \tag{3.6a}$$

$$\beta_{1t} = b_0 + b_1 h_{t|t-1} \tag{3.6b}$$

$$h_{t|t-1}^2 = \sigma^2(1 - \delta - \gamma) + \delta \epsilon_{t-1}^2 + \gamma h_{t-1|t-2}^2 \quad . \tag{3.6c}$$

This approach constitutes a simplified version of the model in Demos (2002), who generalizes the GARCH-M framework of Engle et al. (1987) to the case of stochastic volatility and time-varying coefficients. In (3.6a) to (3.6c), volatility is non-stochastic which eliminates identification issues and drastically alleviates estimation. The model is flexible enough, however, to serve our purpose, i.e. allowing for state-dependent effects. Moreover, in empirical applications, the parsimonious GARCH(1,1) has often been sufficient to control for conditional heteroskedasticity. The use of the standard deviation in (3.6b) has some dampening effect on extreme volatility spikes. We maximize the likelihood function under the assumption of normally distributed shocks. Since the normality assumption is often too restrictive for financial time series data, we rely on quasi-maximum likelihood and obtain robust standard errors as in Bollerslev and Wooldridge (1992).

We use (3.6a) to (3.6c) to test Hypotheses 1 and 2 as follows. We run a series of regressions for several $s_t^{(\tau)}$. Hypothesis 1 is tested by checking whether β_{1t} is positive and increasing in τ for all t. Hypothesis 2 would be supported if $\beta_{1t} > 0$ for all t and $b_1 > 0$ since this would reflect that the impact of D_t increases in risk aversion. Finally, we note that compared to (3.4), where the overall impact of D_t on $s_t^{(\tau)}$ is measured by β_1, the analogue of the total effect in the ADL-GARCH-M model is given by $[\beta_{1t} + \psi(1)] \cdot [1 - \phi(1)]^{-1}$.

[5]In principle, this model can be generalized such that the coefficients of lagged values of D_t in the polynomial $\psi(L)$ are also allowed to vary over time. The specification in (3.6a)–(3.6c), however, already implies the long-run effect, given by $[\beta_{1t} + \psi(1)] \cdot [1 - \phi(1)]^{-1}$, to be time-varying. Moreover, in the empirical application below, lagged values of D_t are found insignificant.

3.4 Data: Yield Spreads and the Maturity Structure of Debt

Since the end of 1997 the Deutsche Bundesbank publishes daily observations of constant maturity yield series. The main empirical analysis below starts at 1/1/1998 and ends at 31/12/2007. We initially cut off data from 2008 onwards to put the focus primarily on the years before the financial crisis. This allows for a meaningful comparison of our results to those of Greenwood and Vayanos (2012), where the crisis is also excluded.[6]

Throughout the remainder of the paper, we will refer to the 6-month rate, the shortest rate available from the Bundesbank data bank, as the short-rate. To provide an overview of the effects of relative supply of longer-term debt on spreads along the maturity spectrum, we consider several maturities of longer-term rates, namely $\tau = 3, 4, 5, 7$ and 10 years. Term spreads are then calculated as the difference between the longer-term rates and the short-rate and are denoted by $s_t^{(3)}$, $s_t^{(4)}$, $s_t^{(5)}$, $s_t^{(7)}$ and $s_t^{(10)}$. Table 3.1 provides some descriptive statistics. On average, spreads are positive and are increasing and more volatile for larger τ.

The issuance of German bonds is executed by the Finanzagentur GmbH. The bonds can be classified in those listed at the stock exchange and those not listed. Since the yield data is based only on traded debt, we ensure consistency by also using solely listed bonds to measure debt supply.[7] The bonds include Federal Treasury notes (maturities ranging from 6 months to 2 years), Five-year Federal notes (maturity of 5 years) and Federal bonds (predominately with a maturity of 10 years, but also with 30 years). Traded debt should still provide a reasonably precise indication of the maturity structure of total German government debt, since from 1998 onwards the fraction of non-traded debt out of total debt decreased quickly and steadily from about 10% to less than 2% (see column 3 of

[6]The extreme increase in interest rate spreads in the course of the financial crisis requires certain adjustment of our model. Results from the extended sample ending at 31/12/2012 are presented and discussed in detail in Appendix 3.A.

[7]Non-traded debt includes Federal Treasury financing paper and Federal savings notes of type A and B. These bonds have maturities similar to listed bonds.

Table 3.2). Over the historical course of debt accumulation, a continuous maturity spectrum of bonds became available at any given point in time. This is particularly true for maturities up to 10 years.

Table 3.1: Spreads and Debt - Descriptive Statistics

	mean	$\hat{\sigma}$	min	max
spreads				
$s_t^{(3)}$	0.457	0.371	-0.350	1.410
$s_t^{(4)}$	0.623	0.449	-0.330	1.700
$s_t^{(5)}$	0.770	0.512	-0.270	1.900
$s_t^{(7)}$	0.971	0.588	-0.150	2.120
$s_t^{(10)}$	1.288	0.707	-0.020	2.530
debt				
D_t	45.351	2.328	36.761	51.589

Notes: This table reports descriptive statistics of the term spreads and the relative supply of longer-term debt. All statistics are measured in percent and calculated from daily observations over the sample 1/1/1998 to 31/12/2007.

Following Greenwood and Vayanos (2012), relative supply of longer-term debt is defined as debt that is to be paid within a certain period in the future divided by the total value of debt. Total debt at t refers to the sum of all principals and coupon payments due until the very last bond is matured. The average maturity of German debt is found to be around 5 years throughout the whole sample period (see Table 3.2). Therefore, we set relative supply of longer-term debt as equal to the fraction that is to be paid in 5 years hence and label it by D_t.[8] Correspondingly, any payments to be made within the coming 5 years are interpreted as shorter-term debt.

[8]Greenwood and Vayanos (2012) define D_t as the fraction of debt that is to be paid in 10 years hence. Particularly in more recent times, however, they also report longer average maturities for US debt of around 7 years.

A time series of D_t is not readily available. Therefore, this paper generates a new data set of relative debt supply. The data required to construct the debt variable are obtained from Bloomberg. According to information directly provided by the Finanzagentur GmbH, the amount of traded debt more than doubled from €438.2 billion on 12/31/1998 to €909.22 billion on 12/31/2007. This matches closely with the data available from Bloomberg, allowing to trace back, on average, 99% of traded debt.

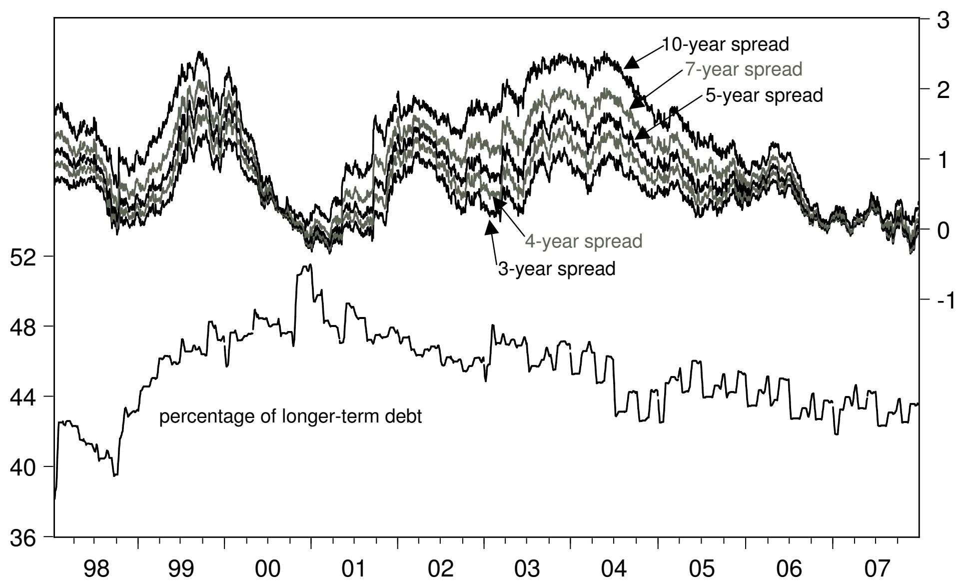
Table 3.2: The Maturity Structure of German Government Debt

end of the year	total debt	$\frac{\text{traded debt}}{\text{total debt}}$	# of bonds	$\bar{\tau}$	percentage of debt due within τ years					
					1	3	5	7	10	30
1998	478.9	89.8	76	4.6	13.2	37.5	56.7	66.7	84.1	100
1999	708.3	93.9	72	4.6	15.1	36.9	56.7	66.7	83.2	100
2000	715.6	94.7	67	4.9	12.2	32.6	48.4	64.5	78.5	100
2001	697.3	96.0	62	4.7	14.6	34.6	52.3	60.3	82.3	100
2002	719.4	97.3	60	4.7	15.5	36.3	53.6	63.5	84.6	100
2003	760.4	98.2	58	4.7	14.5	35.0	53.0	67.0	80.7	100
2004	803.0	98.5	55	4.8	15.7	37.3	55.6	64.1	82.2	100
2005	872.6	98.6	54	4.6	15.7	37.9	54.7	66.3	80.0	100
2006	902.0	98.5	53	4.6	16.4	39.6	56.3	65.7	81.4	100
2007	922.0	98.6	54	4.7	17.2	39.2	56.4	66.9	80.1	100

Notes: This table reports descriptive statistics of the maturity structure of German government debt. Total debt equals the value of all outstanding bonds, i.e. the sum of listed and non-listed bonds. Debt is measured in € billion. $\bar{\tau}$ refers to the average maturity of debt measured in years. Column 2 and 3 are based on data directly provided by the Finanzagentur GmbH. The rest of the statistics is based on data obtained from Bloomberg.

The debt variable is generated as follows. For any bond, we observe the outstanding amount denoted in euro, the issue date, the number of days left until maturity, the principal, the coupon and the coupon frequency. This information allows us to track each bonds' payment flow over its lifetime, i.e. coupon and principal payments. As can be observed in the last row of Table 3.1, longer-term debt roughly varies between a good third and one half, and averages about 45% with a standard deviation of 2.3 percentage points. Both relative debt supply and the term spreads are shown in Figure 3.4.

Figure 3.4: Interest Rate Spreads and Relative Supply of Longer-Term Bonds



3.5 Empirical Results

3.5.1 Static Regressions

We begin the empirical analysis by running the static regression given in (3.4) by OLS. According to Hypothesis 1, the slope coefficient β_1 of this regression should be positive and increasing in τ. Our regression results for the German data can be

Table 3.3: Spreads and Debt - Static Regressions

$$s_t^\tau = \beta_0 + \beta_1 D_t + u_t$$

				Greenwood and Vayanos (2012)	
$s_t^{(\tau)}$	$\widehat{\beta}_1$	R^2	DW	$\widehat{\beta}_1$	R^2
$s_t^{(3)}$	0.003	$0.4 \cdot 10^{-3}$	0.012	0.025^{***}	0.055
	[1.093]			[2.564]	
$s_t^{(4)}$	0.011^{***}	0.003	0.010	0.034^{***}	0.062
	[2.809]			[2.742]	
$s_t^{(5)}$	0.016^{***}	0.005	0.008	0.040^{***}	0.065
	[3.742]			[2.799]	
$s_t^{(7)}$	0.021^{***}	0.007	0.006	$--$	$--$
	[4.265]				
$s_t^{(10)}$	0.030^{***}	0.009	0.005	$0.077^{***\dagger}$	0.097
	[4.913]			[3.677]	

Notes: This table reports results from static regressions. Columns 2 – 4 refer to the results obtained from daily German data over the period 1/1/1998 – 12/31/2007. The numbers is brackets denote t-values and DW refers to the Durbin-Watson statistic. The last two columns show the results reported in Greenwood and Vayanos (2012) which are based on monthly US data over the sample June 1952 – December 2005 and robust standard errors following Newey-West (1987). Estimates for $s_t^{(7)}$ and $s_t^{(10)}$ are not provided and the value indicated by † stems from a regression with $s_t^{(20)}$. *** denotes significance at the 1% level.

found in Table 3.3 together with the results reported by Greenwood and Vayanos (2012), which they obtained from US data in the same specification.

The first major result is that our estimates are positive and increasing in τ. Moreover, apart from the case of $s_t^{(3)}$, the t-values in brackets document that the coefficients are highly significant. Therefore, the static model appears to provide strong evidence in favor of Hypothesis 1, i.e. term spreads widen when the relative supply of longer-term debt increases. Compared to the results for the US data in the last two columns of Table 3.3, our point estimates are consistently lower but still of similar magnitude. The R^2's are considerably higher in the US case. At least in part, however, this could be attributed to the lower (monthly) data frequency.

In fact, the extremely low R^2's of our regressions indicate that almost no variation of the spreads is explained by the relative supply of longer-term debt. Moreover, the Durbin-Watson (DW) statistics in column 4 of Table 3.3 are startling. In all regressions the DW statistics are close to zero, which means that very high first order autocorrelation is present in the residuals. This raises serious concerns about whether the inference in model (3.4) is sound.

3.5.2 Dynamic Regressions and State-Dependent Coefficients

We continue the empirical analysis in two steps. In order to control for the strong autocorrelation present in the static regressions, we estimate the dynamic model (3.5), with two lags of the dependent variable.[9] Thereafter, we estimate the ADL-GARCH-M model (3.6) to test for state-dependent effects. Results are summarized in Table 3.4.

A comparison between columns 2 in Tables 3.3 and 3.4 shows that, once the serial correlation is taken into account, t-values decrease considerably. In fact, the relation between debt supply and spreads vanishes in the dynamic model. At the 5% level, none of the estimated coefficients is significant anymore. Hence, the static regression results were spurious. The Lagrange multiplier statistics LM(10) and corresponding p-values in column 3 indicate that there is no autocorrelation up to order 10.[10] This suggests that the inference in the dynamic model is sound.

[9] The choice to include two lags is based on residual autocorrelation tests. Lags of the independent variable were also considered but found insignificant.

[10] Since the DW statistic tests only for first order autocorrelation and is also biased toward 2 when a lagged dependent variable is included in the regression, we conduct LM tests instead.

Table 3.4: Spreads and Debt - Dynamic Regressions and State-Dependent Coefficients

dynamic regression with constant coefficients	dynamic regression with state-dependent coefficients
$s_t^{(\tau)} = \beta_0 + \beta_1 D_t + \phi_1 s_{t-1}^{(\tau)} + \phi_2 s_{t-2}^{(\tau)} + \epsilon_t$	$s_t^{(\tau)} = \beta_0 + \beta_{1t} D_t + \phi_1 s_{t-1}^{(\tau)} + \phi_2 s_{t-2}^{(\tau)} + \epsilon_t$ $\beta_{1t} = b_0 + b_1 h_{t\|t-1}$ $h_{t\|t-1}^2 = \sigma^2(1-\delta-\gamma) + \delta\epsilon_{t-1}^2 + \gamma h_{t-1\|t-2}^2$

$s_t^{(\tau)}$	$\hat{\beta}_1$	LM(10)	$\hat{b}_1$	R^2_{static}	total effect $\hat{\beta}_{1t}\cdot(1-\hat{\phi}_1-\hat{\phi}_2)^{-1}$ if risk aversion is *low*	*mean*	*high*	impact of $\hat{\sigma}$-change in D_t
$s_t^{(3)}$	$0.5\cdot 10^{-3}$	1.508	0.007^{***}	0.103	0.021	0.034	0.060	5 bp
	[1.474]	(0.130)	[3.080]					14 bp
$s_t^{(4)}$	$0.6\cdot 10^{-3}$	1.340	0.008^{***}	0.200	0.026	0.043	0.078	6 bp
	[1.652]	(0.203)	[3.598]					18 bp
$s_t^{(5)}$	$0.7\cdot 10^{-3}$	1.287	0.008^{***}	0.258	0.033	0.052	0.099	8 bp
	[1.755]	(0.232)	[3.310]					22 bp
$s_t^{(7)}$	$0.7\cdot 10^{-3}$	0.841	0.008^{***}	0.333	0.042	0.066	0.126	10 bp
	[1.830]	(0.589)	[2.982]					29 bp
$s_t^{(10)}$	$0.8\cdot 10^{-3}$	0.577	0.007^{***}	0.412	0.049	0.080	0.141	11 bp
	[1.801]	(0.834)	[2.727]					33 bp

Notes: This table reports results from dynamic regressions with constant coefficients (columns 2 and 3) and state-dependent coefficients (columns 4 – 9). Numbers in brackets show t-values based on robust standard errors from Bollerslev and Wooldridge (1992). *** and ** indicate significance at the 1% and 5% level. LM(10) denotes the F-statistic of a Lagrange multiplier test for autocorrelation up to order 10. The corresponding p-values are given in parentheses. R^2_{static} refers to the R^2 from a *static* regression with state-dependent coefficients and reflects the explained variation that is not simply due to the inclusion of lags. Columns 6 – 8 document the total effect of D_t on $s_t^{(\tau)}$, depending on the state of risk aversion. The last column presents the total impact under low and high risk aversion of a one standard deviation shock in D_t on $s_t^{(\tau)}$, measured in basis points (bp).

We now turn to the estimates obtained from the ADL-GARCH-M model shown in columns 4 – 9 of Table 3.4.[11] Most strikingly, the coefficient b_1, that governs the state-dependency in the relation between $s_t^{(\tau)}$ and D_t, is positive and highly significant, which is in line with Hypothesis 2. That is, the time-varying coefficient specification reveals that there actually is a relation after all, which would have remained undiscovered in the constant coefficient model.[12]

In order to allow for a comparison of our results to those from the US data, we calculate the total effect of a change in debt supply on spreads. While the overall impact in the static model is simply given by the slope coefficient β_1, in the dynamic model the response accumulates due to the lags. Hence, in the ADL-GARCH-M specifications the total effect is given by $\hat{\beta}_{1t} \cdot (1 - \hat{\phi}_1 - \hat{\phi}_2)^{-1}$. Since $\hat{\beta}_{1t}$ depends on $h_{t|t-1}$, columns 6 – 8 of Table 3.4 report the values of $\hat{\beta}_{1t}$ for the minimum, mean and maximum value of the conditional standard deviation. First, we compare column 7 of Table 3.4 with column 5 of Table 3.3. Under the *mean* level of risk aversion, we find a total effect that is almost the same as the one for the US. Moreover, the fact that the values are increasing from 0.034 to 0.080 not only supports Hypothesis 1 but also Hypothesis 2. The results under *low* and *high* risk aversion highlight the relevance of the state-dependency. In times of high risk aversion, the response of the term spread to changes in debt supply is up to 3 times as high as in times of low risk aversion.

As to the explained variation, we consider the statistic R^2_{static} reported in Table 3.4. In order to allow for a meaningful comparison of the R^2's in our ADL-GARCH-M regressions with the R^2's from the US data, we excluded the lagged values. This is because all R^2's in the dynamic specification are almost 1 due to the autoregressive components. Therefore, the statistic R^2_{static} refers to a static regression with a state-dependent slope coefficient. The values are fairly large,

[11] As in the dynamic regression with constant coefficients, including two lags is based on residual autocorrelation tests. Lags of the independent variable were also considered but found insignificant. Autocorrelation and heteroskedasticity specification tests as well as additional estimation results can be found in Appendix 3.C.

[12] The constant term b_0 was found to be insignificant without exception. This is in line with the result that β_1 is not significant in the constant coefficient ADL model. If there were constant effects, one would expect to see them also in the standard ADL specification.

ranging from about 10% to more than 40%. Accordingly, relative supply of longer term debt - *if adjusted by risk aversion* - has substantial explanatory power for term spreads.

The last column of Table 3.4 illustrates the economic relevance of the parameter estimates. We calculated the long-run reaction of $s_t^{(\tau)}$ to a one standard deviation shock in D_t. Since the standard deviation of D_t is about 2.3 percentage points, such a shock would roughly equal a shift of €21 billion of debt from shorter to longer maturities. The exact widening of the term spread would depend on τ and the level of risk aversion. From the shorter end to the middle of the maturity spectrum, we see a reaction between 5 and 22 basis points (bp). At the longer end, we observe an impact between 10 and 33 bp.

3.6 Conclusion

Building on Modigliani and Sutch (1966), recent approaches in the term structure literature elaborate on the role of preferred-habitat investors (Vayanos and Vila 2009, Greenwood and Vayanos 2010, Greenwood and Vayanos 2012, Guibaud et al. 2013). Bond prices are understood to be determined by the supply of bonds through the government and by the demand for bonds through preferred-habitat investors and arbitrageurs. The models predict that an increase in the relative supply of longer-term debt should drive up interest rate spreads. Preferred-habitat effects are, however, supposed to be more pronounced when risk aversion of arbitrageurs is high and their participation in the bond market is limited.

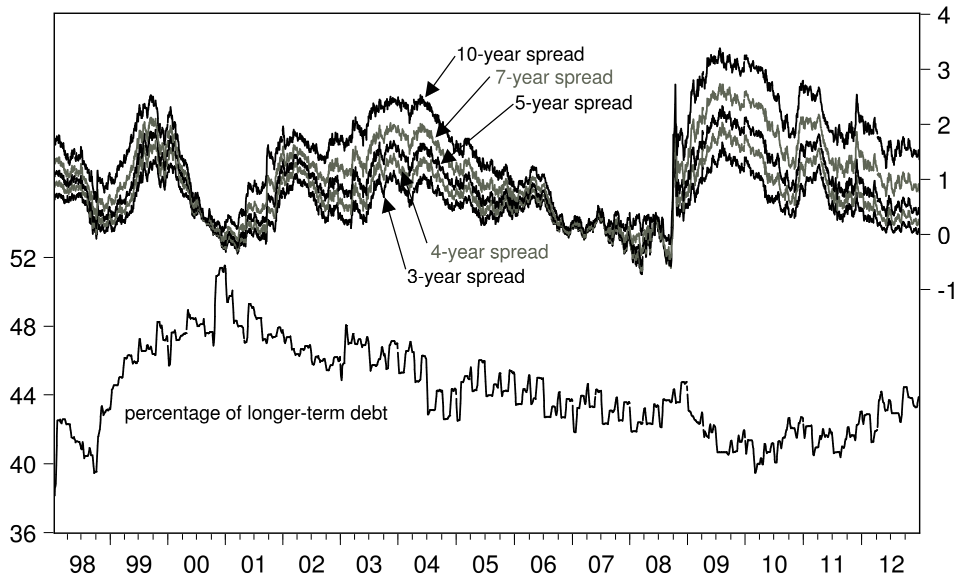
This paper argues that the degree of risk aversion is central to the empirical analysis of the preferred-habitat theory. We propose an econometric framework that is flexible enough to account for changing risk aversion by allowing for state-dependent coefficients. Moreover, our methodology takes the strong autocorrelation, present in term spreads and debt supply, into account. Formally, we introduce an ADL-GARCH-M where the conditional standard deviation proxies the degree of risk aversion and governs the state-dependency of the coefficients in the mean equation. We apply the model to a new data set of daily observations of relative supply of longer-term debt in Germany.

Our results suggests that there is a significantly positive relation between yield spreads and the relative supply of longer-term debt, which crucially depends on the state of risk aversion. In line with the model predictions, the impact of debt supply on term spreads is stronger for larger differences in maturities between long-term and short-term rates. For all analyzed spreads it holds that the reaction to changes in debt supply is approximately three times larger in times of high risk aversion than in those of low risk aversion. The responses of spreads to a one standard deviation increase in debt supply varies between 5 and 33 basis points. Moreover, a static regression with constant coefficients would substantially underestimate the effect of debt supply on the term spread.

Due to the decisive role of risk aversion that we document empirically, our results suggest that explicit theoretical modeling of time-varying preference parameters may provide valuable new insights into the role that is played by preferred-habitat investors in bond markets. The policy implication of preferred-habitat models is that a change in the maturity structure of government debt alters bond yields. On the basis of German bond data, this paper supports that view. There is, however, a crucial reservation: the effect may only be of sufficient economic relevance in relatively turbulent times characterized by high volatility and high risk aversion. Hence, bond purchasing programs as the Outright Monetary Transactions of the European Central Bank should be most effective in crisis times.

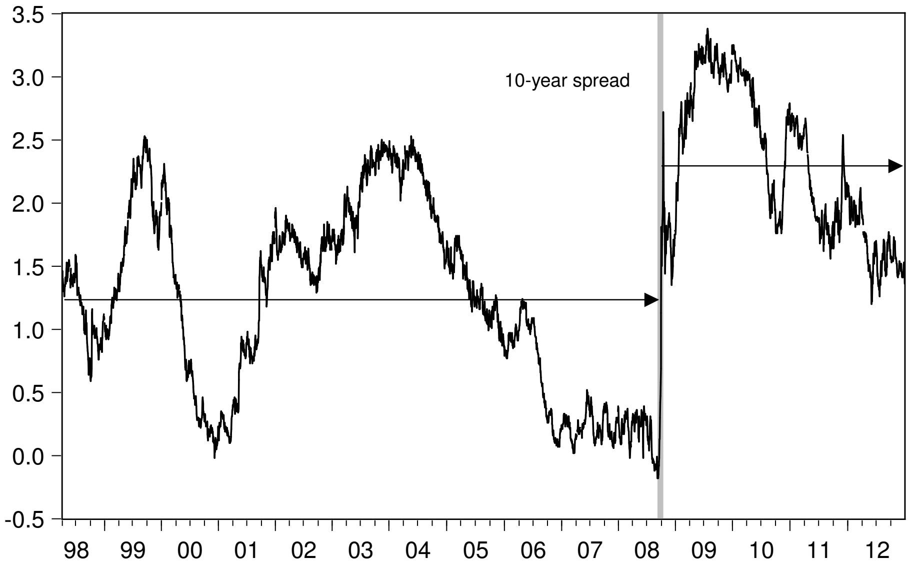
3.A Results from the Extended Sample

Figure 3.5: Yield Spreads and Relative Supply of Longer-Term Bonds



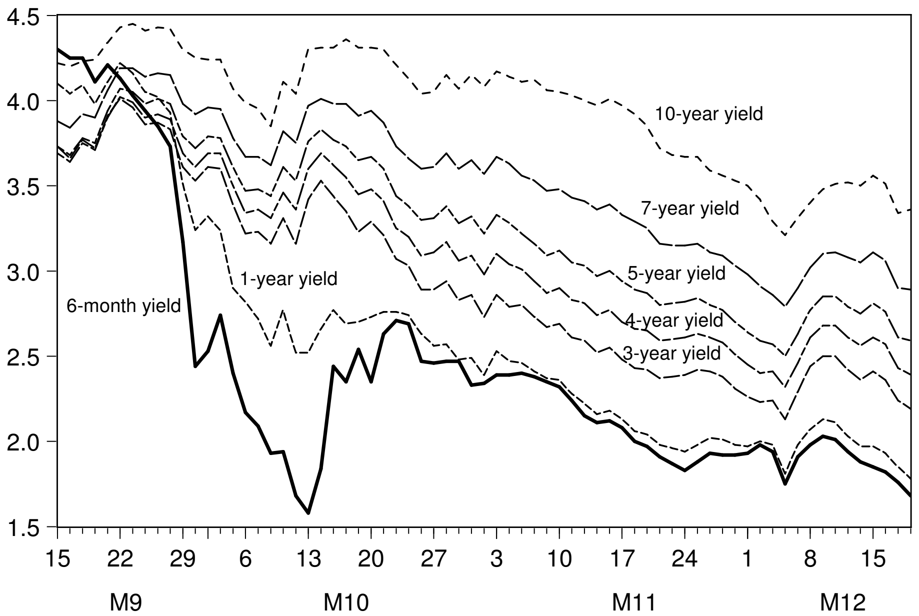
During the ongoing financial crisis, bond yields show extraordinary developments, especially at the short end of the maturity spectrum. This makes the recent sample period particularly interesting to examine. At the same time, however, the econometric methodology may require some adjustment towards this new regime. In the present case, we modify our framework in two respects. Firstly, we include a shift dummy, $\tilde{\beta}_0 d$, in our ADL-GARCH-M regressions that allows for a structural break in the constant term at the Lehman crash. Secondly, due to the extraordinary movements associated with flight-to-safety effects at the very short end of the yield curve during the end of 2008, we use the 1-year yield as short-rate over the extended sample period. When we simply ignore the enormous shift in spreads being clearly visible in Figure 3.5, our results from Section 3.5.2 do not hold.

Figure 3.6: 10-Year Yield – 6-Month Yield



In order to visualize the structural break more clearly, Figure 3.6 shows only $s_t^{(10)}$ as a representative example. Comparing the empirical means before and after September 2008, we observe an increase from about 1.25% to 2.25%. We take this structural change into account by modeling it as structural break in the unconditional mean. To recognize the difference between the 6-month yield, which was considered as the short-rate in the main empirical analysis, and the 1-year yield, see Figure 3.7. Compared to the other yields, the 6-month rate drops drastically from a level that exceeds the ones of longer-term yields before the Lehman crash, to a remarkably low level with a trough at 1.5%. This suggests that movements at the very short end of the yield curve are largely driven by extreme events, such as extensive usage of German short-term bonds as a safe haven for banks to temporarily place funds. The 1-year yield already shows a pattern that seems much more closely linked to longer-term bonds. Therefore, we replaced the 6-month rate by the 1-year rate over the extended sample period.

Figure 3.7: Bond Yields After the Lehman Crash 2008



Results in the extended sample are given in Table 3.5. As in the shorter sample period, we find $\hat{b}_1$ to be positive, significant and increasing in τ. Furthermore, in view of columns 4 and 6, the minimum and maximum values of β_{1t} show that β_{1t} is positive for all t. Therefore, the results provide supportive evidence for Hypotheses 1 and 2. Compared to the shorter sample period, the variation in β_{1t} due to changing risk aversion has increased. This reflects the strong increase in our risk aversion proxy during the crisis.

Table 3.5: Spreads and Debt - Dynamic Regressions and State-Dependent Coefficients

1/1/1998 – 12/31/2012

dynamic regression with

state-dependent coefficients

$$s_t^{(\tau)} = \beta_0 + \beta_{1t} D_t + \phi_1 s_{t-1}^{(\tau)} + \phi_2 s_{t-2}^{(\tau)} + \tilde{\beta}_0 d + \epsilon_t$$

$$\beta_{1t} = b_0 + b_1 h_{t|t-1}$$

$$h_{t|t-1}^2 = \sigma^2(1-\delta-\gamma) + \delta\epsilon_{t-1}^2 + \gamma h_{t-1|t-2}^2$$

			total effect $\hat{\beta}_{1t} \cdot (1-\hat{\phi}_1-\hat{\phi}_2)^{-1}$ if risk aversion is			
$s_t^{(\tau)}$	$\hat{b}_1$	R^2_{static}	*low*	*mean*	*high*	impact of $\hat{\sigma}$-change in D_t
$s_t^{(3)}$	0.003***	0.030	0.015	0.032	0.089	4 bp
	[2.679]					22 bp
$s_t^{(4)}$	0.004***	0.080	0.029	0.047	0.123	7 bp
	[3.318]					31 bp
$s_t^{(5)}$	0.005***	0.133	0.034	0.058	0.145	9 bp
	[3.394]					37 bp
$s_t^{(7)}$	0.005***	0.237	0.044	0.074	0.173	11 bp
	[3.190]					44 bp
$s_t^{(10)}$	0.005***	0.337	0.051	0.085	0.182	13 bp
	[3.001]					46 bp

Notes: This table reports results from dynamic regressions with state-dependent coefficients. Numbers in brackets show t-values based on robust standard errors from Bollerslev and Wooldridge (1992). *** and ** indicate significance at the 1% and 5% level. R^2_{static} refers to the R^2 from a *static* regression with state-dependent coefficients. Columns 4 – 6 document the total effect of D_t on $s_t^{(\tau)}$, depending on the state of risk aversion. The last column presents the total impact under low and high risk aversion of a one standard deviation shock in D_t on $s_t^{(\tau)}$, measured in basis points (bp).

3.B Results for the Limiting Case: Non-Stationarity of Term Spreads and Debt Supply

As a further robustness check, we consider the extreme case of a unit root in term spreads and also in the relative supply of longer-term debt. Even though both variables are clearly bounded from an economic point of view, and hence should be stationary, $I(1)$ processes may empirically provide the best approximation of the data generating process. Whether this is the case, however, often remains unclear. The outcome of unit root tests can, among other things, depend crucially on the null hypothesis specified by the researcher.

We apply two tests: the GLS-ADF test of Elliott et al. (1996) with the null of a unit root and the KPSS test of Kwiatkowski et al. (1992) with the null of a stationary process, see Table 3.6. In the extended sample we use the test of Zivot and Andrews (1992) (ZA) instead of the GLS-ADF test. The ZA test allows for an endogenous structural break in the unconditional mean, which is motivated by Figures 3.5 and 3.6. It should be noted that the ZA test finds the break at the Lehman crash, just as we have specified in our ADL-GARCH-M regressions in Appendix 3.A.

As can be seen from Table 3.6, regardless of the sample, both tests fail to reject the null – non-stationarity or stationarity – at any conventional level. If we followed the GLS-ADF and ZV test results, we would conclude that all variables contain a stochastic trend. In that case, the following equations would represent a more convenient representation of the ADL-GARCH-M model.

$$\Delta s_t^{(\tau)} = \alpha(c + s_{t-1}^{(\tau)} + \beta_{1t} D_{t-1}) + \omega(L)\Delta D_t + \kappa(L)\Delta s_{t-1}^{(\tau)} + \epsilon_t \tag{3.7a}$$

$$\beta_t = b_0 + b_1 h_{t|t-1} \tag{3.7b}$$

$$h_{t|t-1}^2 = \sigma^2(1 - \delta - \gamma) + \delta\epsilon_{t-1}^2 + \gamma h_{t-1|t-2}^2 \quad . \tag{3.7c}$$

Table 3.6: Spreads and Debt - Unit Root Tests

	1/1/1998 – 12/31/2007		1/1/1998 – 12/31/2012	
	GLS-ADF	KPSS	ZA	KPSS
spreads				
$s_t^{(3)}$	-1.469	0.172	-3.762	0.200
$s_t^{(4)}$	-1.184	0.184	-3.631	0.218
$s_t^{(5)}$	-0.986	0.199	-3.687	0.233
$s_t^{(7)}$	-0.761	0.239	-3.638	0.242
$s_t^{(10)}$	-0.587	0.312	-3.700	0.244
debt				
D_t	-0.646	0.195	-0.866	0.377

Notes: This table reports unit test results of the GLS-ADF test (Elliott et al. 1996), the KPSS test (Kwiatkowski et al. 1992) and the ZA test (Zivot and Andrews 1992). To the variable D_t the GLS-ADF test is applied in both samples. *** and ** indicate rejection at the 1% and 5% level.

The framework in (3.7a) – (3.7c) constitutes an error correction model with a time-varying cointegrating vector. Accordingly, the parameter β_{1t} now has a different interpretation than in the ADL model, i.e. it represents the total effect. The test statistic for a cointegration relation is given by the t-value of α. It is not immediately clear, however, which critical values should be applied. Banerjee et al. (1998) provide critical values for single equation error correction models with a constant cointegrating vector. For the present model, where we have time-varying coefficients, a simulation experiment showed that the critical values of Banerjee et al. (1998) are also valid. The following steps indicate the design of our simulation.

Step 1. Draw two random samples of size $N = 2539$ (equal to the number of observations in the present analysis) from a standard normal distribution. Denote these shocks by $\xi_{s,t}$ and $\xi_{d,t}$.

Step 2. Generate data under the null of no cointegration. The term spread $s_t^{(\tau)}$ follows an integrated autoregressive process of order 2 with GARCH(1,1) errors driven by $h_{t|t-1}\,\xi_{s,t}$. The relative supply of longer-term debt D_t follows an integrated autoregressive process of order 2 driven by $\xi_{d,t}$. Set the parameters equal to those obtained from estimating the model under the null.

Step 3. Estimate model (3.7a) – (3.7c) via ML (BHHH algorithm) using the generated series of spread and debt supply. Save the t-value of $\hat{\alpha}$ based on robust standard errors following Bollerslev and Wooldridge (1992).

Step 4. Repeat *Step 1* to *Step 3* 25,000 times.

Step 5. Calculate the 5.00 and 10.00 percentiles from the distribution of the t-value of $\hat{\alpha}$.

For $s_t^{(3)}$, $s_t^{(4)}$, $s_t^{(5)}$, $s_t^{(7)}$ and $s_t^{(10)}$, the point estimates of the long-run multiplier remain unchanged. The t-values of the α's are -3.385, -3.617, -3.406, -3.095 and -2.732. These values can be compared to the critical valules in Banerjee et al. (1998). The 10% and 5% quantiles are given by -2.89 and -3.19. Hence, apart from $s_t^{(10)}$, the results even survive the $I(1)$ case, at least at the 10% significance level. We conclude that a significant state-dependent relation in levels between term spreads and the relative supply of longer-term debt exists. Whether this is a cointegration relation or a relation between two stationary variables is not the pivotal question since neither the interpretation of the estimates nor the test decisions in the inference hinge on that distinction.

Table 3.7: State-Dependent Coefficients: Estimation Results and Specification Tests

dynamic regression with state-dependent coefficients

$$s_t^{(\tau)} = \beta_0 + \beta_{1t} D_t + \phi_1 s_{t-1}^{(\tau)} + \phi_2 s_{t-2}^{(\tau)} + \epsilon_t$$

$$\beta_{1t} = b_0 + b_1 h_{t|t-1}$$

$$h_{t|t-1}^2 = \sigma^2(1-\delta-\gamma) + \delta\epsilon_{t-1}^2 + \gamma h_{t-1|t-2}^2$$

$s_t^{(\tau)}$	$\hat{\beta}_0$	$\hat{b}_1$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\delta}$	$\hat{\gamma}$	Q(5)	Q(10)	LM(5)	LM(10)
$s_t^{(3)}$	-0.009	0.007	0.824	0.168	0.053	0.933	4.072	14.485	1.242	1.224
	[-2.581]	[3.080]	[39.007]	[7.954]	[5.333]	[74.913]	(0.539)	(0.152)	(0.286)	(0.270)
$s_t^{(4)}$	-0.010	0.008	0.824	0.167	0.055	0.933	2.026	11.980	1.038	0.973
	[-2.933]	[3.598]	[40.045]	[8.119]	[5.883]	[83.130]	(0.846)	(0.286)	(0.393)	(0.464)
$s_t^{(5)}$	-0.011	0.008	0.813	0.179	0.052	0.935	1.002	9.086	1.154	0.727
	[-2.759]	[3.310]	[40.893]	[9.087]	[6.998]	[105.878]	(0.962)	(0.524)	(0.329)	(0.706)
$s_t^{(7)}$	-0.012	0.008	0.808	0.186	0.042	0.947	1.040	5.472	1.1115	0.646
	[-2.805]	[2.982]	[38.715]	[8.894]	[5.750]	[110.266]	(0.956)	(0.857)	(0.350)	(0.775)
$s_t^{(10)}$	-0.012	0.007	0.804	0.190	0.027	0.967	1.271	3.706	0.782	0.553
	[-2.570]	[2.727]	[38.181]	[8.983]	[5.623]	[166.068]	(0.938)	(0.960)	(0.563)	(0.853)

Notes: This table reports estimation results and specification tests from dynamic regressions with state-dependent coefficients over the sample 1/1/1998 to 12/31/2007. Numbers in brackets show t-values based on robust standard errors following Bollerslev and Wooldridge (1992). The coefficient b_0 was found insignificant and was set to zero in order to gain efficiency. Q(5) and Q(10) represent Q-statistics for remaining autocorrelation in ϵ_t up to order 5 and 10 respectively. LM(5) and LM(10) denote F-statistics of Lagrange multiplier test for remaining GARCH effects up to order 5 and 10 respectively. Corresponding p-values in both specification tests are given in parentheses.

4 The Signal of Volatility

4.1 Introduction

The present study examines the economic interpretation of volatility in financial markets and proposes a new and flexible econometric approach. Firstly, we pinpoint two fundamental understandings of volatility that have emerged from the financial literature during the last decades. On the one hand, the fact that prices vary is interpreted as a sign of *information* flow. On the other hand, high variability is often seen as a mirror image of pronounced *uncertainty* in the market. Both views suggest volatility-dependent stock market interaction, albeit in different directions, and we aim at shedding light on the inherent ambivalence. In a simple economic framework, we show that higher volatility in one market should lead to higher (lower) reactions in another market if volatility reflects information (uncertainty). To the best of our knowledge, these two views of volatility have never been explicitly contrasted and empirically examined. Secondly, we propose a strategy to infer the dominating signal of return variability from the data: we analyze different reactions of investors to observed returns, depending on the prevailing level of volatility. As our econometric framework, we introduce a simultaneous time-varying coefficient model, where time variation is a function of ARCH-type variances. The analysis is based on daily data of major stock indexes from the Americas, Australia and the Asian region.

Let us first provide some background concerning the two signals of volatility we put up for discussion and review some literature we see connected to our line of reasoning. From one point of view, volatility is often associated with uncertainty or risk. Considering the global financial crisis for instance, future market developments are highly uncertain. In the public discussion, the image of labile and disoriented financial markets prevails. Intuitively, the extensive stock market volatility is often interpreted as the reflection of this uncertainty. In the present study this concept of volatility shall be summarized as the *uncertainty hypothesis*.

Regarding the pricing of assets, it seems natural that investors expect to be compensated for bearing uncertainty in their portfolios. In fact, in academia the understanding of volatility as risk long plays an important role with a prominent example given by the μ-σ-utility function and the CAPM. Originating from Engle et al. (1987), financial econometricians translated this idea into the variance-in-mean model (see also French et al. 1987, Bali and Engle 2010 and the references therein). Another example for volatility proxying uncertainty is given by interactions between output or inflation uncertainty and the conditional means of these variables (e.g. Grier and Perry 2000). In a further strand of literature, numerous studies analyze how uncertainty about exchange rate movements affects trade volume and foreign direct investment, e.g. Cushman (1985), Chowdhury (1993) and Kiyota and Urata (2004). For instance, volatility might negatively impact the size of trade flows if exchange rate uncertainty renders trade less profitable for risk averse agents.

On the other hand, we will refer to the view of volatility being a measure of information flow intensity as the *information hypothesis*. Some representatives of the literature who elaborate on the volatility-information link are Clark (1973), Epps and Epps (1976), Ross (1989) and Fleming et al. (1998). Overall, the idea is that no motivation for further trading would exist in a situation where all prices have settled at their equilibrium values. Thus, volatility would be zero in absence of relevant news. If, however, additional information becomes available, price adjustments will generate fluctuations until a new equilibrium is reached. Of course, in reality, shocks are too frequent to allow conventional asset prices to ever settle at some constant consensus value, and perception and handling of information both represent more complicated processes than assumed in stylized model economies. Nonetheless, the line of reasoning exemplifies how volatility is connected to information arrival.

The information content of price movements is normally not observable. This is likely to be one of the main reasons why information flow was associated with volatility in the first place. By the same token, a strand of literature examined trading volume as an observable variable that is at least partly driven by the information arrival process; see Tauchen and Pitts (1983), Harris (1987), Lamoureux and Lastrapes (1990b), Foster and Viswanathan (1993, 1995), Gagnon and Karolyi (2009). Certainly, volume cannot *explain* volatility, in the sense of an exogenous variable. Instead, both are affected simultaneously by the latent information process. Moreover, many trades are unlikely to be linked to information arrival, such as in the cases of liquidity management (e.g. Andersen 1996), strategic trading under asymmetric information (e.g. Kyle 1985) or differences of opinions on the interpretation of signals (e.g. Kim and Verrecchia 1991). Attempts have been made to proxy information arrival directly by, for example, central bank decisions, macroeconomic news or firm-specific announcements. For studies of corresponding volatility effects, see e.g. Andersen and Bollerslev (1998), Kalev et al. (2004) or Goeij and Marquering (2006). Nonetheless, even if important insights into news effects could be gained, such direct observable measures cannot represent more than a fraction of the universe of information arriving in financial markets. Above all, they hardly capture private information, which is a major factor behind volatility (French and Roll 1986).

Our distinct hypotheses serve to fix ideas concerning the character of volatility. Naturally, they are not mutually exclusive. Rather, exploring the "signal of volatility" amounts to asking which effect predominates. In fact, this calls for a mechanism connecting the latent variables information and uncertainty to a measure that is estimable from the data. In the present approach, we propose letting the reaction of market participants decide the character of volatility instead of leaving this task up to the econometrician. Specifically, we make use of the intensity by which shocks feed into actual market prices, thereby connecting a high intensity to high information content, as further explained below. However, given a single observed time series, identifying the size of the shocks themselves (i.e., volatility) and the size of their impact on the price separately, proves evidently impossible.

We approach this problem by extending the information set to the multivariate case. In particular, we examine the intensity of spillover between two different markets. Logically, while shocks can be identified in the "source" market, transmission intensity is measured in the "target" market. In case observed price changes in the source market are interpreted as highly informative (uncertain) signals by the target market, the latter will incorporate a relatively large (small) fraction of the innovation into its own price. We illustrate this principle in a stylized model economy, based on signal extraction by rational agents. Overall, high volatility in the target market associated with high spillover intensity would support the information hypothesis, while evidence for the uncertainty hypothesis would follow from an inverse linkage.

Econometrically, we measure this nonlinear effect in a time-varying coefficient model governed by the (autoregressive) conditional variance of the source market, i.e., we utilize time variation in volatility to identify its impact on transmission intensity. Such an empirical strategy has not yet been considered in the literature. Our concept does not aim at explaining the mere fact that markets are interconnected, e.g. by trade, policy coordination or common shocks. Rather, we exploit the existing interaction for estimating the spillover intensity and its link to volatility. Furthermore, the a priori division into "source" and "target" markets is an artificial one. In reality, once one introduces spillover effects, one must take a stance on how to resolve endogeneity. Our model set-up will generally allow for bi-directional transmission between the US and the second country of interest. Identification is achieved by making use of the heteroskedasticity in the data, which can be exploited to uniquely pin down the structure of simultaneous systems; compare Sentana and Fiorentini (2001) or Rigobon (2003). Therefore, both the direction and the size of spillovers can be determined empirically. These considerations on simultaneity apply to markets with overlapping trading hours, like in the Americas. For models of the US and the major Asian or Australian stock indexes, the spillover direction is given by the sequence of time, since these markets trade with substantial time shifts. Consequently, identification problems are alleviated in this setting.

Our first major result is that in all countries under investigation spillover intensity significantly depends on volatility. As regards the information content of volatility, our results tell that it crucially depends on the combination of "sender" and "receiver" of volatility signals. For industrial countries, the information hypothesis holds. As for emerging economies, however, the uncertainty hypothesis prevails in their relations to the US.

The rest of the paper proceeds as follows. The next section presents a stylized model of stock market returns and derives the testable hypotheses. Section 4.3 introduces the econometric model and discusses identification issues and the estimation procedure. Section 4.4 applies the methodology to daily returns of major stock indexes from the Americas, Australia and the Asian region. The last section concludes.

4.2 Volatility Signals in a Stylized Model Economy

4.2.1 The Market Participant: Signal Extraction Problem

First we illustrate the idea of the signal of volatility in a stylized model economy. This should help fix ideas on how stock market interaction could depend on return variability. Moreover, the nature of this interdependence should reveal the character of volatility, i.e., it should indicate whether volatility in one market means information or uncertainty (noise) to the other. A prominent model from the literature, which can be used for this purpose, was considered by King and Wadhwani (1990). We adopt this framework to demonstrate that in a signal extraction context, the prevailing character of volatility can be identified from the optimal reaction of investors to observed returns.

For the present purpose, it is sufficient to consider two stock markets where price changes are associated with the arrival of relevant information and with noise, i.e., uncertainty. The first consists of two parts: directly observed information and a reaction to information that is not fully observed in that market but only in the other:

$$y_{1t} = \iota_{1t} + \alpha_{12}\mathbb{E}[\iota_{2t}|I_{1t}] + \nu_{1t} \tag{4.1}$$

$$y_{2t} = \alpha_{21}\mathbb{E}[\iota_{1t}|I_{2t}] + \iota_{2t} + \nu_{2t} \quad . \tag{4.2}$$

Stock returns are given by y_t, information is denoted by ι_t, ν_t refers to noise and $\mathbb{E}[\cdot|I_{jt}]$ represents the expectations operator conditional on the information observed in market j at time t. The model reflects the usually positive correlation of international stock returns, i.e. $\alpha_{12} \geq 0$, $\alpha_{21} \geq 0$.

When investors form expectations, say in market 1, they face a simple signal extraction problem, since all they can observe from market 2 is the contemporaneous price change. In order to extract the signal from the part of the price movement in market 2 that is not simply due to information in market 1, agents in market 1 have to find β_1 in

$$\mathbb{E}[\iota_{2t}|I_{1t}] = \beta_1(y_{2t} - \alpha_{21}\mathbb{E}[\iota_{1t}|I_{2t}]) \quad . \tag{4.3}$$

The solution to (4.3) is given by the minimum-variance estimator:

$$\beta_1 = \frac{\mathrm{Var}[\iota_{2t}]}{\mathrm{Var}[\iota_{2t}] + \mathrm{Var}[\nu_{2t}]} \quad . \tag{4.4}$$

Evidently, β_1 becomes time varying, i.e., β_{1t}, in case volatility of either ι_{2t} or ν_{2t} changes over time.

Of course, agents in market 2 follow an analogous rationale. Using (4.3) and (4.4) to substitute for the conditional expectations in (4.1) and (4.2) yields the following simultaneous equations system of stock returns:

$$y_{1t} = A_{12t}y_{2t} + \epsilon_{1t} \tag{4.5}$$

$$y_{2t} = A_{21t}y_{1t} + \epsilon_{2t} \quad , \tag{4.6}$$

where the spillover coefficients are given by $A_{12t} = \alpha_{12}\beta_{1t}$ and $A_{21t} = \alpha_{21}\beta_{2t}$. The shocks result as $\epsilon_{1t} = (1-\alpha_{12}\alpha_{21}\beta_{1t}\beta_{2t})(\iota_{1t}+\nu_{1t})$ and $\epsilon_{2t} = (1-\alpha_{12}\alpha_{21}\beta_{1t}\beta_{2t})(\iota_{2t}+\nu_{2t})$.

In our application, we will choose the US as the first country and switch between several other stock markets in y_2. Logically, the model will change according to the choice of the second country. In addition to the second equation, this concerns also the first one, (4.1). Apart from the spillover, the partitioning of the return shock into information and noise, and thus also β and A, depend on the perspective of the second country. In order to keep the notation simple, we write down model (4.1)-(4.2) only for a given set of countries.

4.2.2 The Econometrician: Testable Hypotheses

The model (4.1)-(4.2) cannot be estimated directly since all four shocks are unobservable. Otherwise, we could simply estimate their variances to see which volatility effect dominates. However, we have shown that under the assumption that the model in (4.5) and (4.6) is identified, one could exploit the time variation in the spillovers in order to measure volatility signals. Following the reasoning from above, the contemporaneous impact from one market to the other depends on the variances of both signal (information) and noise (uncertainty). The econometrician can approach the problem of measuring volatility signals by estimating the variance of ϵ_t, i.e. the entire shocks to the returns. Taking the typical time-varying nature of financial time series volatility into account, we denote the conditional variance of ϵ_t by $\mathrm{Var}[\epsilon_t|I_{t-1}] = h_t$ and let the spillover coefficients depend on the variances by

$$A_{ijt} = f_{ij}(h_{jt}) \qquad i,j = 1,2 \quad \text{and} \quad i \neq j \; . \tag{4.7}$$

As can be seen in (4.4), beta would be constant if the variation, i.e. the rate of change in $\mathrm{Var}[\iota_{jt}|I_{t-1}]$ and $\mathrm{Var}[\nu_{jt}|I_{t-1}]$, was exactly identical. Assume, for instance, that $\frac{\partial f_{ij}}{\partial h_{jt}} > 0$, so that $\mathrm{Var}[\iota_{jt}|I_{t-1}]$ dominates the dynamics of market volatility in the sense that its rate of change is higher than the one of $\mathrm{Var}[\nu_{jt}|I_{t-1}]$. This would favor the information hypothesis. On the contrary, $\frac{\partial f_{ij}}{\partial h_{jt}} < 0$ would represent evidence for the uncertainty hypothesis. In sum, examining the time variation in spillover strength can provide us with decisive information on which of the shocks contributes more to the volatility dynamics.

While the theoretical model in the previous section serves as a motivation, we argue that empirically the functional form of $f(\cdot)$ over the whole value range is not known a priori. As discussed in detail in the next section, allow for flexibility by approximating $f(\cdot)$ on an empirical basis. So far, we summarize the following two testable hypotheses:

Information Hypothesis:
The spillover intensity A_{ijt} in (4.5) and (4.6) depends *positively* on the level of volatility in the respective other market, i.e., $\frac{\partial A_{ijt}}{\partial h_{jt}} > 0$.

Uncertainty Hypothesis:
The spillover intensity A_{ijt} in (4.5) and (4.6) depends *negatively* on the level of volatility in the respective other market, i.e., $\frac{\partial A_{ijt}}{\partial h_{jt}} < 0$.

4.3 Empirical Approach: Measuring Investors Reaction to Observed Returns

4.3.1 Simultaneous Model and Identification

In order to explore the signal of volatility, we first discuss our simultaneous model setup. The considered stock returns are collected in the n-dimensional vector y_t. The data generating process is approximated by the following simultaneous system:

$$Ay_t = \mu_t + \varepsilon_t \ , \tag{4.8}$$

where μ_t represents a vector of predictable components such as lags or a constant term and ε_t is a n-dimensional vector of structural innovations. The contemporaneous impacts are included in matrix A with diagonal elements normalized to one. It is these effects that model the spillovers between returns in the current setting and that we will allow to depend on volatility later on. Common shocks will be accommodated by allowing for correlation of ε_t, as explained below.

The simultaneous specification (4.8) is not meant to take a stance on *fundamental* causality, in the sense that an impulse say in market j is necessarily the true causal origin of a spillover to market i. Of course, one can think of idiosyncratic events in market j affecting market i, based on economic linkages or psychological effects. However, an impulse in market j may well be initiated by some information that is equally relevant for market i, where investors observe the signal from j. Then it would evidently be the third-party origin of the information, and not market j itself, which would underlie the impact on market i. In summary, spillovers characterize *signals* in one stock index that are incorporated by other markets, but not necessarily based on actual bivariate causality.

Statistically, model (4.8) as it stands is not identified: In the matrix A with a normalized diagonal, $n(n-1)$ simultaneous impacts have to be estimated, whereas the covariance matrix of the reduced-form residuals $A^{-1}\varepsilon_t$ delivers only $n(n-1)/2$ determining equations due to its symmetry. However, as for instance Sentana and Fiorentini (2001) and Rigobon (2003) show, unobservable factor structures like (4.8) become unique if heteroskedasticity is present in the stochastic components. The idea is that, although breaks in the structural variances introduce additional unknowns (i.e., the variances in the new regime), they shift the whole covariance matrix in the reduced form, from which available information (i.e., variances and covariances) is doubled. Time-varying volatility is a common feature of financial variables, often modeled as ARCH-type processes. Indeed, the approach of Sentana and Fiorentini (2001) subsumes the case of regime switches just as other forms of heteroskedasticity such as ARCH. Here, we follow Weber (2010), who specifies multivariate EGARCH processes for the structural shocks.

Formalizing the model setup, first denote the conditional variances of the elements in ε_t by

$$\mathrm{Var}(\varepsilon_{jt}|\Omega_{t-1}) = h_{jt}^2 \qquad j = 1, \ldots, n \ , \tag{4.9}$$

where Ω_{t-1} stands for the whole set of available information at time $t-1$.

Furthermore, denote the standardized innovations by

$$\tilde{\varepsilon}_{jt} = \varepsilon_{jt}/h_{jt} \qquad j = 1, \ldots, n \ . \tag{4.10}$$

EGARCH(1,1)-processes are then given by

$$\log h_{jt}^2 = c_j + g_j \log h_{jt-1}^2 + d_j(|\tilde{\varepsilon}_{jt-1}| - \sqrt{2/\pi}) + f_j \tilde{\varepsilon}_{jt-1} \qquad j = 1, \ldots, n \ , \quad (4.11)$$

where c_j, g_j, d_j and f_j represent the coefficients. The term $\sqrt{2/\pi}$ serves to demean the absolute shock. In addition, going beyond the pure magnitude of shocks, the signed $\tilde{\varepsilon}_t$ introduce asymmetric volatility effects. The logarithmic formulation ensures positive variances without relying on parametric restrictions.

Common shocks are taken into account via the structural constant conditional correlation (SCCC) approach of Weber (2010). The advantage of the SCCC model is to relax the uncorrelatedness assumption for structural shocks on the one hand but to keep up the identification of the simultaneous model achieved through heteroskedasticity on the other. The covariances of structural shocks are recovered by the CCC specification

$$\mathrm{Cov}(\varepsilon_{it}, \varepsilon_{jt} | I_{t-1}) = h_{ijt} = \rho_{ij} h_{it} h_{jt} \qquad i \neq j \ , \qquad (4.12)$$

where ρ_{ij} denotes the correlation between the ith and the jth innovation.[1] This correlation can be thought of as arising from the exposure of variables i and j to unobserved common factors.

For markets with non-overlapping trading hours identification problems are alleviated. The fact that country i is only trading while the stock exchange in country j is closed implies that contemporaneous spillovers do not appear. In our model setup, this amounts to specifying a triangular coefficient matrix A_t. Even though the index t then does not refer to the same time for all variables, we keep the notation for simplicity purposes.

4.3.2 Time-Varying Coefficients

Up to this point, the off-diagonal elements of matrix A in (4.8) imply spillovers between the endogenous variables that are proportional to the size of shocks with

[1] We also considered the structural *dynamic* conditional correlation (SDCC) approach. However, empirical evidence for time variation could not be found.

proportionality factors constant over time. While this represents the standard in simultaneous systems, the current research question requires a more complex specification. Therefore, we develop a framework that combines the heteroscedastic structural model introduced above with a time-varying spillover specification. In order to discriminate between the information and uncertainty hypotheses, we allow the transmission intensity to depend on source market volatility as derived in section 4.2.2.

Strictly speaking, A is substituted by A_t in (4.8). The elements A_{ijt}, $i \neq j$, denote the coefficients of transmission from variable j to i at time t. As a parsimonious functional form, consider the linear specification of (4.7):

$$A_{ijt} = a_{ij} + b_{ij} h_{jt} \ , \tag{4.13}$$

for all i, j. Here, the conditional standard deviation h_{jt} serves as the transition variable. Since A_t stands on the left hand side, negative values represent positive transmission. Therefore, a_{ij} is expected to be smaller than zero. Accordingly, a one-unit increase in source market volatility decreases spillover intensity by b_{ij}. Hence, from the above discussion it follows that $b_{ij} < 0$ would favor the information hypothesis, whereas prevalence of the uncertainty hypothesis requires $b_{ij} > 0$. Alternatively, $b_{ij} = 0$ would bring us back to the case of constant parameters.

We note that this specification can be compared to the GARCH-in-mean model, where returns are explained by their own conditional variances. In our approach, the variance series is also employed for an interaction effect with the level. However, we allow the spillover in one mean equation to depend on the conditional variance of another return.

No case can be made, a priori, that the transition function (4.13), i.e., the volatility effect on spillover intensity, is necessarily linear. While the advantage lies in parametric parsimony, the exact functional form of (4.7) should be determined on an empirical basis. For instance, let us assume a situation with $a < 0$ and evidence for the uncertainty hypothesis, say $b > 0$. At a certain point, a linear transition function could approach a negative correlation between markets (i.e., with a positive left-hand-side coefficient). Since such a constellation appears rather implausible, the transition effect is likely to exhibit dampening non-linearity for high volatility values. Still, if such realizations are rare in the sample, (4.13) might work well as approximation of the transition function (4.7).

As an alternative specification, literature on smooth transition regression (STR) (e.g. Luukkonen et al. 1988) has adopted flexible functions to grasp time variation in coefficients. Specifically, consider

$$A_{ijt} = a_{ij} + \alpha_{ij}/(1 + e^{-\gamma_{ij}(h_{jt}-\beta_{ij})}) \; . \tag{4.14}$$

The exact form of the transition is determined by the logistic function $(1 + e^{-\gamma(h-\beta)})^{-1}$, which is monotonically increasing[2] in h_{jt} and bounded between zero and one. The slope parameter γ indicates the speed or smoothness of transition: as $\gamma \to \infty$, the logistic function approaches the indicator function $I(h_{jt} > c)$, i.e., a single threshold. In contrast, $\gamma = 0$ simply gives the linear case. The parameter β represents the location of the transition. In sum, the STR-based specification lets the data decide about the shape of the volatility effect on spillover size.

Nonlinear functional forms are one way of dealing with large realizations of the conditional standard deviation. Another straightforward option is given by transforming the transition variable. While we use the standard deviation, taking logarithms as in (4.11), for instance, would further dampen extreme volatility spikes. While there is little reason to believe that a "correct" option could be chosen on theoretical grounds, our results proved robust in this respect.

A last comment concerns the testing of statistical significance of the transition variables in the STR setup. Luukkonen et al. (1988) show that straightforward hypotheses like $\alpha_{ij} = 0$ or $\gamma_{ij} = 0$ are inappropriate because of the presence of unidentified nuisance parameters under the null. Instead, for testing purposes the functions are approximated by a Taylor series of a higher order, usually of order three:

$$A_{ijt} = a_{ij} + b_{ij,1}h_{jt} + b_{ij,2}h_{jt}^2 + b_{ij,3}h_{jt}^3 \; . \tag{4.15}$$

[2]We think of volatility effects on transmission strength being monotonous, even if they are not necessarily linear. More involved STR functions should thus not be required.

Here, standard likelihood ratio (LR) principles apply to the hypothesis $b_{ij,1} = b_{ij,2} = b_{ij,3} = 0$. Of course, linearization may adversely affect the power of the test. However, as Skalin (1998) points out, simulation-based techniques would be extremely computationally demanding and bootstrapping does not provide superior size and power properties. Therefore, we will rely on the LR test in the transition model (4.15). Furthermore, if $b_{ij,2} = b_{ij,3} = 0$ but $b_{ij,1} \neq 0$ is found, the transition function can be approximated by the linear specification (4.13). We maximize the likelihood function under the assumption of normally distributed shocks. As the normality assumption is usually too restrictive for financial time series data, we rely on quasi- maximum likelihood.

4.4 Application: The Signal of International Stock Market Volatility

4.4.1 Data

We examine a balanced sample from 1/1/1988 to 12/31/2010 of daily returns on major stock indices from the US (S&P 500) and a second country of interest. From the Americas we choose Canada (S&P/TSX 60), Argentina (TOTMKAR[3]), Brazil (Bovespa Index) and Mexico (IPC) as examples for contemporaneous trading. The markets of Australia (S&P/ASX 50), Japan (Nikkei), Korea (KOSPI) and the Philippines (PSEi) are all located overseas from the US and represent markets with non-overlapping trading hours.

Stock returns are depicted in Figure 4.1. The time variation in volatility appears very pronounced in all series. This is also statistically indicated by significant autocorrelation of squared returns found in preliminary data inspection. The presence of heteroskedasticity is of special importance to our approach, as it allows estimation of volatility effects on spillover intensity.

[3]Due to data availability for Argentina we use the TOTMKAR provided by Datastream instead of the MERVAL, see http://product.datastream.com/navigator/HelpFiles/DatatypeDefinitions/en/3/DSGI_total_market_data.htm.

4.4.2 Specification Tests

The set of equations to be estimated consists of bivariate simultaneous models with conditional heteroskedasticity for the US and a second country of interest. The empirical application starts with specifying the functional form of the transition function by means of likelihood ratio tests. The specification test procedure can be described as follows:

Since stock market trading hours in Canada and the US are exactly the same and those in Argentina, Brazil and Mexico largely coincide with the US, we allow for bi-directional simultaneous effects. Identification is achieved through the SCCC approach. In the Asian region and Australia, stock markets open after those in the US have closed so that identification issues are alleviated due to this chronology. Hence, we only test for the functional form of the transition function in one direction.

Firstly, we test the null of constant coefficients against linearly time-varying coefficients in all countries. Secondly, the null of linear spillover in both directions is tested separately against the alternative of nonlinear (STR) spillover. In view of the third order Taylor approximation this translates into testing two linear restrictions in (4.15) for each case: H_0: $b_{12,2} = b_{12,3} = 0$ and H_0: $b_{21,2} = b_{21,3} = 0$, respectively.

Columns 2 and 3 of Table 4.2 include p-values of LR specification tests corresponding to the null given in the first row. Bold numbers reflect rejection of the respective null. Column 4 shows the final model specification.[4] [5] To mention one example, in the case of the US and Canada (second row), we find evidence in favor of linear spillover on the US (not rejecting the null in column 2) and nonlinear spillover on Canada (rejecting the null in column 3).

[4]In two cases we do not follow the outcome of the specification tests, namely the Argentinian and Brazilian spillover on the US. Even though statistically nonlinear effects are indicated by the p-values, we restrict the spillover to zero. A closer analysis of these two cases revealed that the smooth transition function actually serves as a dummy to capture only very few outliers at the beginning of our sample while the spillover on the US is otherwise constant and close to zero (between 1% and 2%).

[5]Estimation results from constant coefficients and linearly time-varying coefficient models are provided in Appendix 4.A.

Figure 4.1: Daily Stock Returns on (a) S&P 500, (b) S&P/TSX 60, (c) TOTMKAR, (d) Bovespa Index, (e) IPC, (f) S&P/ASX 50, (g) Nikkei, (h) KOSPI and (i) PSEi

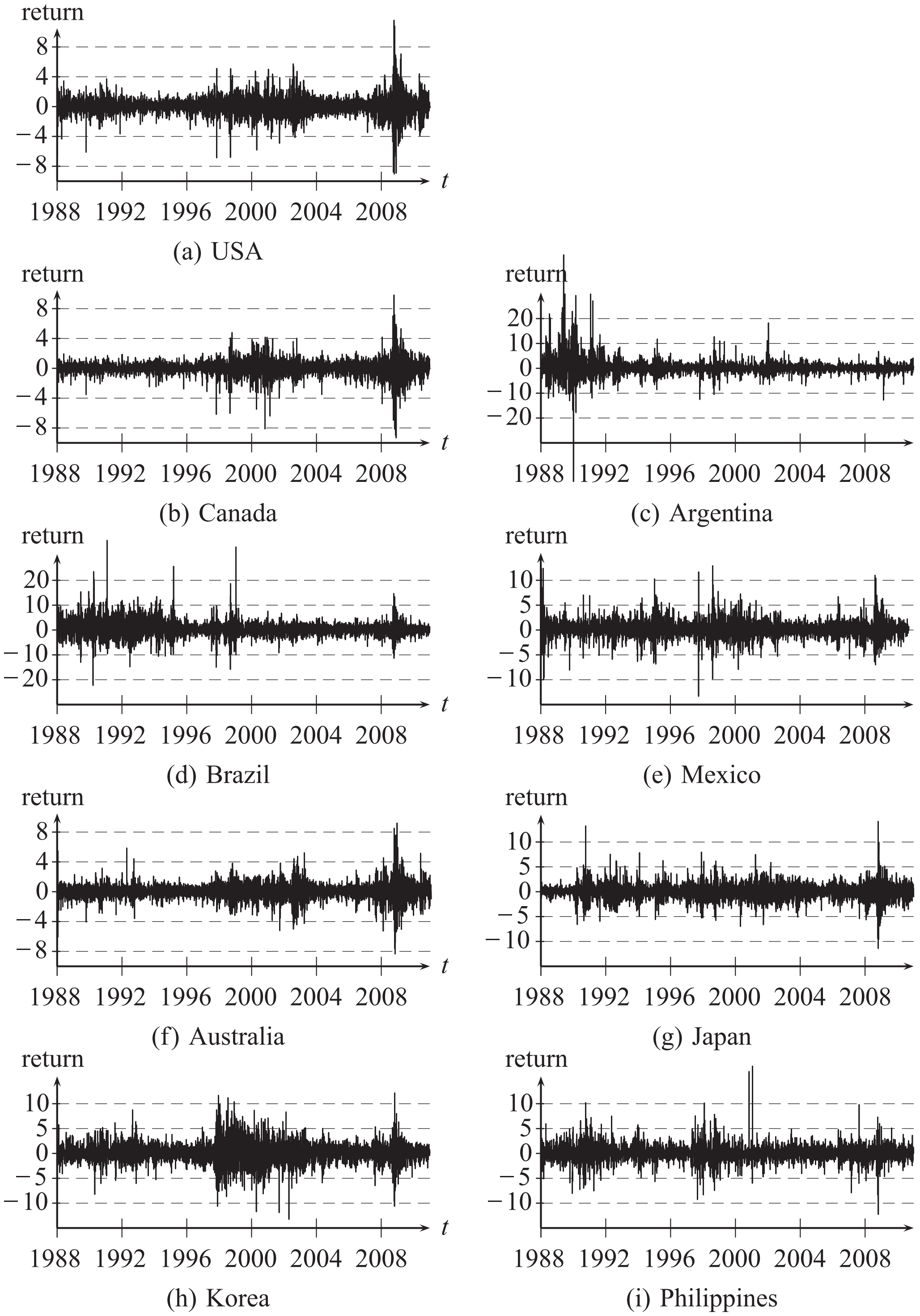


Table 4.1: Specification Tests and Estimation Results

X	H_0: linear on US H_1: STR on US p-values for df = 2	H_0: linear on X H_1: STR on X p-values for df = 2	final model specification	coefficient estimates				signal of volatility
Canada	0.54	**0.05**	linear on US	$a_{12} = 0$	$b_{12} = -0.45$			information
			STR on Canada	$a_{21} = 9.16$	$\alpha_{21} = -9.46$	$\gamma_{21} = 9.47$	$\beta_{21} = -0.19$	information
Australia	–	0.25	no spillover on US					–
			linear on Australia	$a_{21} = -0.35$	$b_{21} = -0.08$			information
Japan	–	**0.04**	no spillover on US					–
			STR on Japan	$a_{21} = -0.28$	$\alpha_{21} = -0.18$	$\gamma_{21} = 32.51$	$\beta_{21} = 0.34$	information
Korea	–	**0.00**	no spillover on US					–
			STR on Korea	$a_{21} = -0.20$	$\alpha_{21} = -0.22$	$\gamma_{21} = 25.65$	$\beta_{21} = 0.40$	information
Argentina	0.03	**0.00**	no spillover on US					–
			STR on Argentina	$a_{21} = -11.31$	$\alpha_{21} = 10.73$	$\gamma_{21} = 5.74$	$\beta_{21} = -0.40$	uncertainty
Brazil	0.03	**0.00**	no spillover on US					–
			STR on Brazil	$a_{21} = -13.30$	$\alpha_{21} = 12.58$	$\gamma_{21} = 5.58$	$\beta_{21} = -0.56$	uncertainty
Mexico	**0.00**	0.09	STR on US	$a_{12} = 0$	$\alpha_{12} = -0.04$	$\gamma_{12} = 14.81$	$\beta_{12} = 0.64$	information
			linear on Mexico	$a_{21} = -0.78$	$b_{21} = 0.10$			uncertainty
Philippines	–	**0.00**	no spillover on US					–
			STR on Philippines	$a_{21} = -0.45$	$\alpha_{21} = 0.16$	$\gamma_{21} = 22.49$	$\beta_{21} = 0.72$	uncertainty

Notes: Columns 2 and 3 report p-values of likelihood ratio tests of the indicated null hypotheses with degrees of freedom equal to df. Bold numbers reflect the rejection of the null. In Argentina and Brazil, we restricted the spillover on US to zero even though test statistics point to nonlinear spillovers; see also footnote 7, page 13. The final specification of the functional form for the time-varying spillover is found in column 4. Columns 5 to 8 show the estimated coefficients. The last column lists the signal for market i that emerges from volatility in market j. Linear or STR specifications of the transition function refer to $A_{ijt} = a_{ij} + b_{ij}h_{jt}$ and $A_{ijt} = a_{ij} + \alpha_{ij}/(1 + e^{-\gamma_{ij}(h_{jt}-\beta_{ij})})$ of the simultaneous model:

$$\begin{pmatrix} 1 & A_{12t} \\ A_{21t} & 1 \end{pmatrix} y_t = \varepsilon_t \ .$$

During estimation we set μ_t constant, as autocorrelation of returns is mostly very close to zero. Results turn out to be insensitive to the inclusion of lagged terms in (4.8). Furthermore, standardized squared residuals appear free from autocorrelation. Thus, we can be confident that our parsimonious EGARCH(1,1) specification is sufficient to capture the time variation in the volatility series.

4.4.3 Results

The first major result is that we find evidence for time-varying spillover coefficients in all countries under investigation. In particular, LR tests (not presented in Table 4.2) of constant against linearly time-varying spillover result in p-values of 0.00 (Canada), 0.08 (Australia), 0.03 (Mexico), 0.000 (Argentina), 0.02 (Brazil), 0.02 (Japan), 0.000 (Korea) and 0.03 (Philippines).[6] That is, for all countries test results suggest the rejection of constant parameters.

Estimated coefficients are presented in columns 5 to 8 of Table 4.2. The hypothesis favored by our evidence is listed in the last column. The results can be divided into two groups. First, the information hypothesis prevails in Australia, Canada, Japan and Korea as US volatility increases the fraction of US shocks that feed into stock prices of these countries. The same holds for Canadian volatility, signaling information for US traders. Second, Argentinian, Brazilian, Mexican and Philippine stock markets seem to understand US volatility as uncertainty since higher volatility leads to a reduction of spillover intensity in these markets. Considering the opposite direction, we find the information hypothesis to dominate in the US with respect to Mexican volatility. However, the small effect from Mexico on the US is economically of minor importance. The transition functions of these markets are plotted in Figures 4.2 to 4.9 (right hand side) together with the spillover intensities (left hand side). We obtain the following results.

Evidence for the Information Hypotheses in Industrial Economies

- In Canada, the effect of US volatility is quite pronounced, indicated by a steep transition function. This results in a transmission that varies between 10% in times of low and approximately 30% in times of high volatility.

[6] The p-value for Korea refers to a test of constant against non-linear spillover. Testing constant against linear spillover yields a p-value of 0.14.

- The information signaling effect of Canadian volatility is also substantial. It produces an even higher spillover variation on the US but, of course, with a lower mean.

- In Australia the information signaling US volatility leads to spillover intensity between roughly 36% and 43%. The spike towards the end of the sample resulting from high US volatility during the crisis drives up transmission strength to 50%.

- Transition functions in Japan and Korea are both strongly increasing in a range of low volatility. Spillover intensity increases for higher levels of volatility by up to 20 percentage points.

Evidence for the Uncertainty Hypotheses in Emerging Economies

- The transition functions and spillover intensity for Argentina and Brazil are of similar shape. In Argentina, however, transmission strength varies around a lower level (70%) than in Brazil (80%). US volatility strongly reduces spillover intensity and is thus interpreted as signaling uncertainty. In both cases, the variance of domestic shocks is high compared to the US, and also to Australia and Canada. Thus, despite high spillover, domestic shocks represent a major factor of return variation in Argentina and Brazil.

- Analogously, transmission strength takes values between 60% and 76% in Mexico with an average of 73% and US volatility having a negative impact. On the contrary, in the US, Mexican volatility increases spillover. Yet, economically the effect fluctuating between zero and a few percent appears to be of secondary importance.

- The contemporaneous impact from the US on the Philippines equals about 45% during times of low volatility. When volatility approaches 1, spillover strongly decreases and falls below 30%.

Figure 4.2: Spillover and Transition Function for Canada and the US

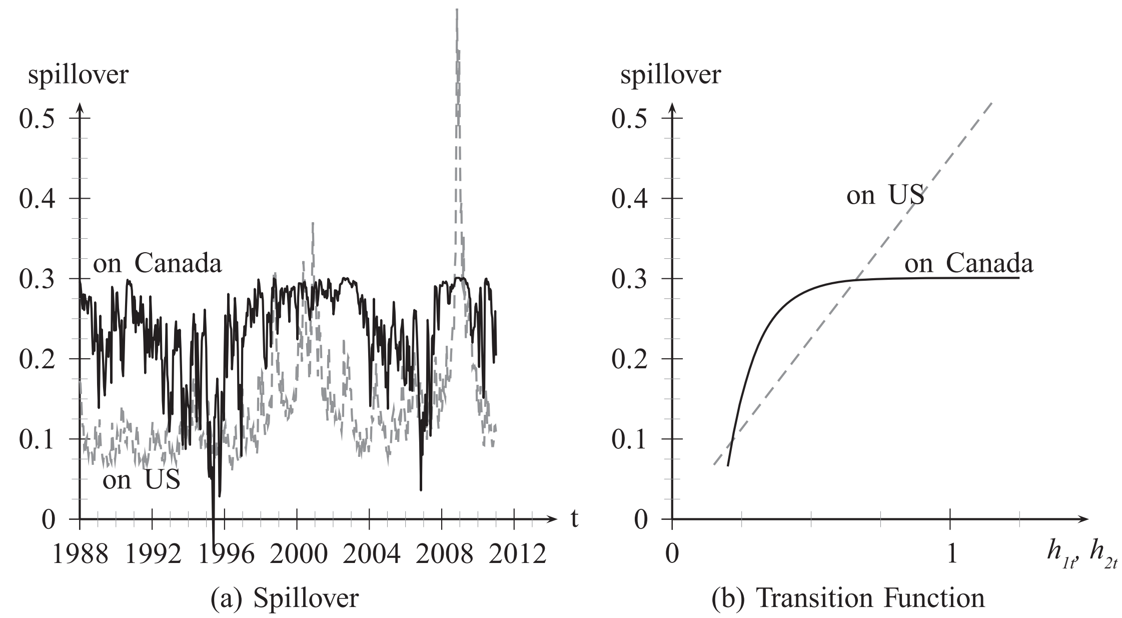


Figure 4.3: Spillover and Transition Function for Australia

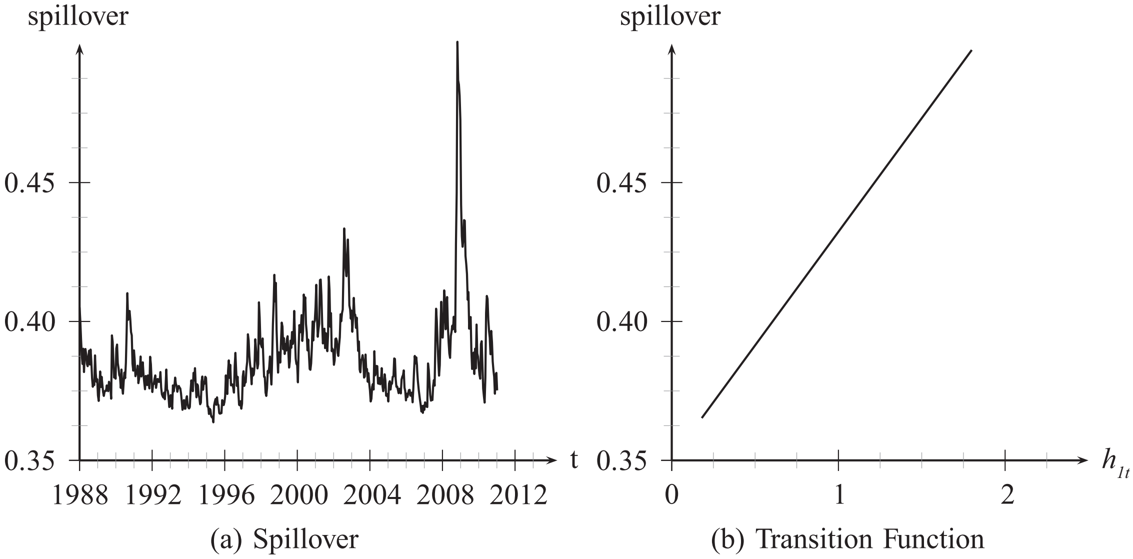
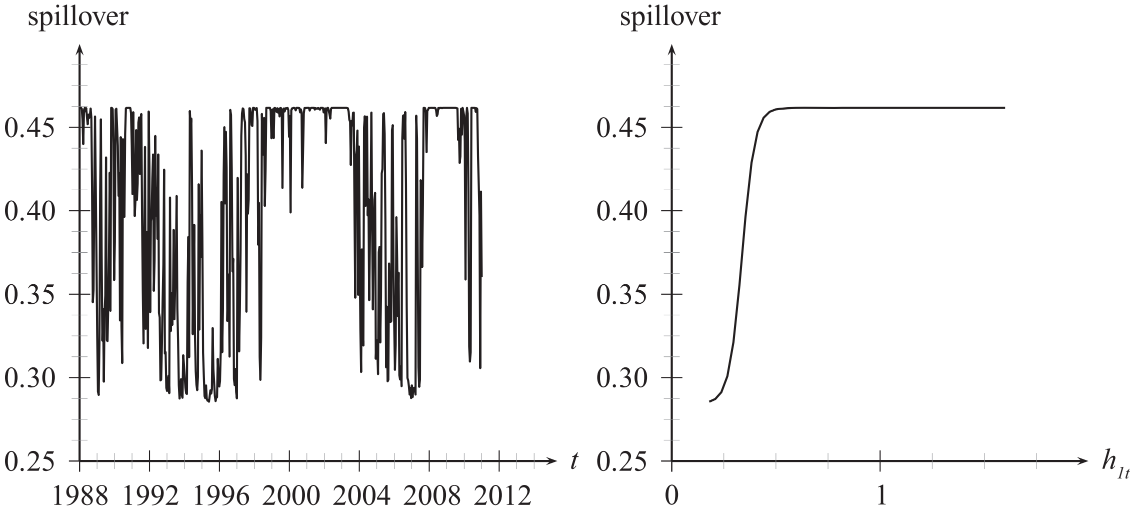
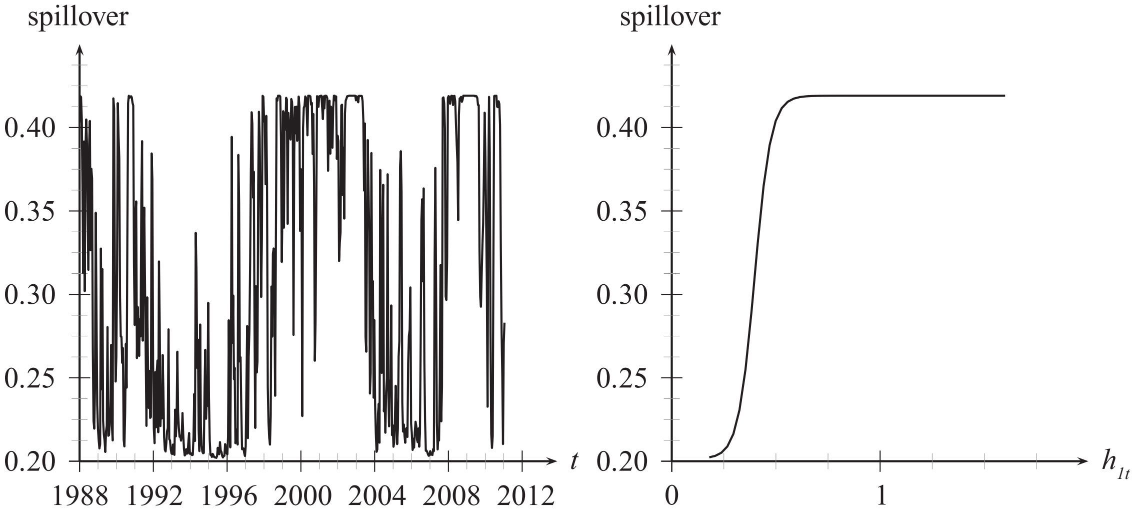


Figure 4.4: Spillover and Transition Function for Japan



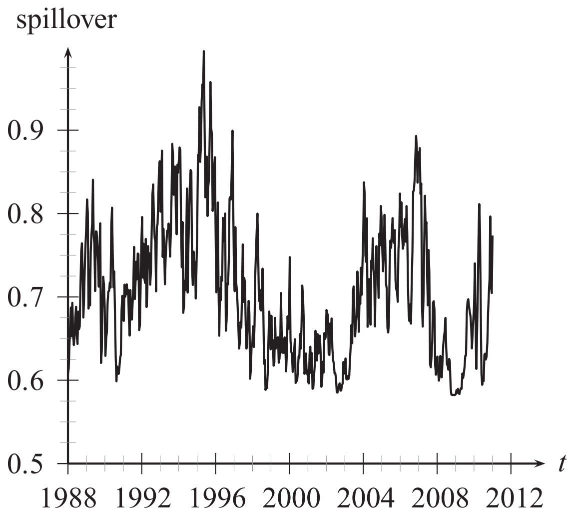
(a) Spillover (b) Transition Function

Figure 4.5: Spillover and Transition Function for Korea

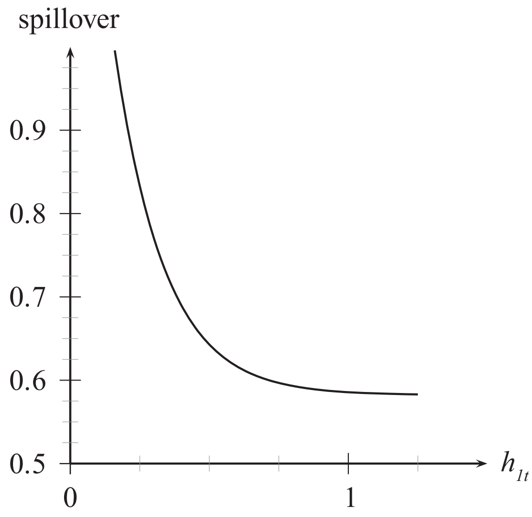


(a) Spillover (b) Transition Function

Figure 4.6: Spillover and Transition Function for Argentina

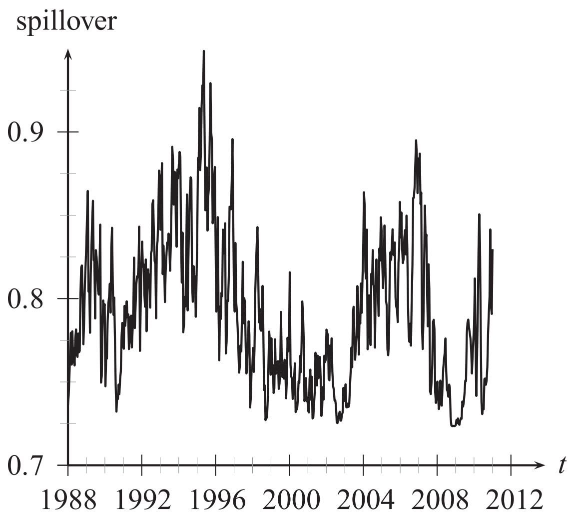


(a) Spillover

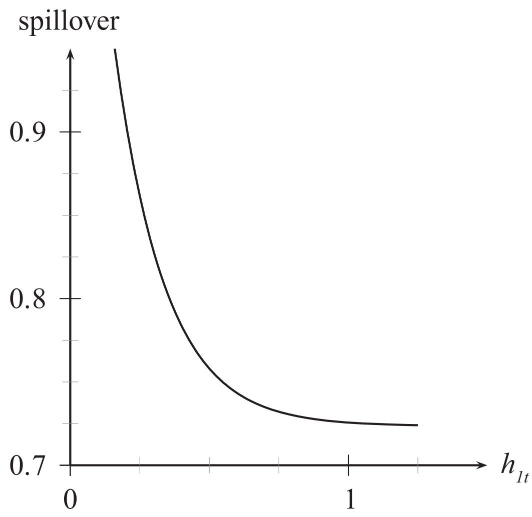


(b) Transition Function

Figure 4.7: Spillover and Transition Function for Brazil



(a) Spillover



(b) Transition Function

Figure 4.8: Spillover and Transition Function for Mexico and the US

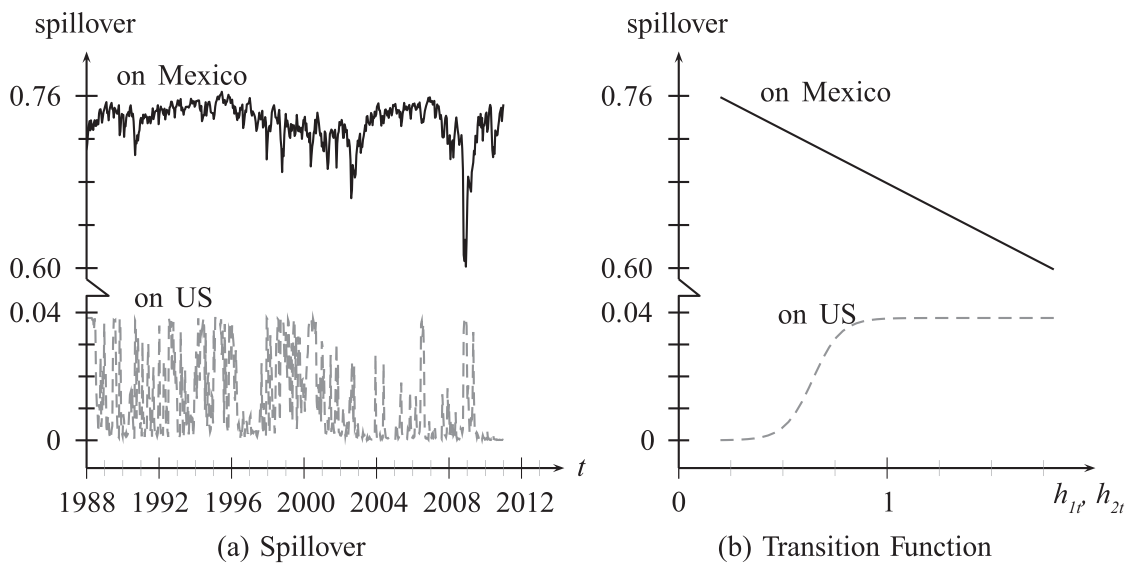
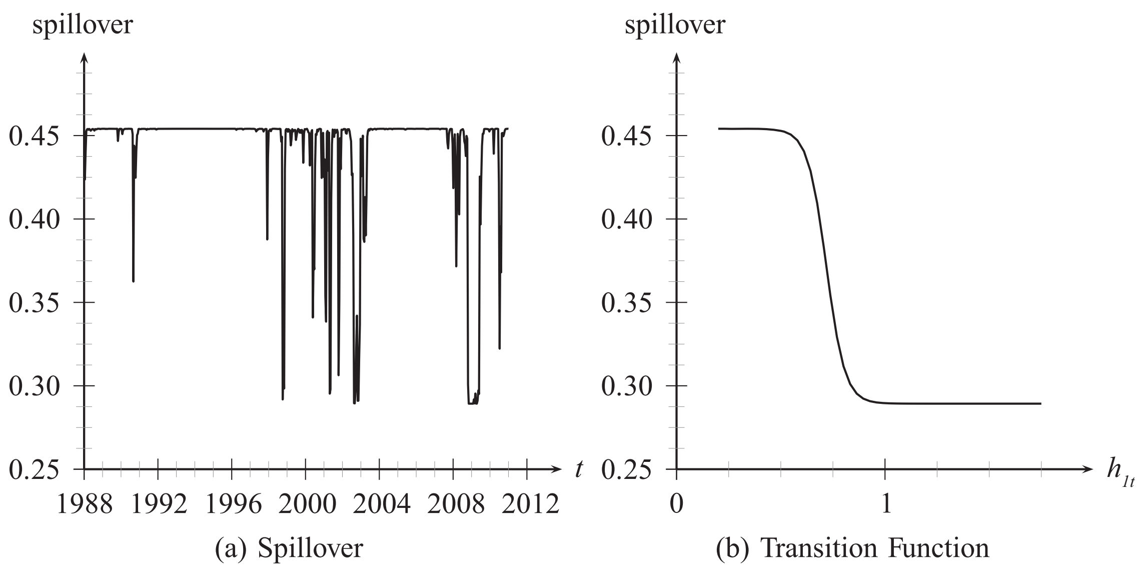


Figure 4.9: Spillover and Transition Function for the Philippines



Interpreting the Stock Market Evidence

Returning to the discussion at the beginning of the paper, the answer to the question whether volatility predominantly signals information or uncertainty is - literally - in the eye of the beholder. On the one hand, identifying shocks in the "source" market and measuring their impact on transmission intensity in the "target" market renders identification and estimation possible. On the other hand, this implies one particular combination of "sender" and "receiver" of volatility signals in each model. The differences in the results across countries show that this combination is crucial. The generally high level of US spillover on the countries under investigation indicates the important role of US stock market developments as a major point of reference. However, even though the "sender of volatility" remains the same in all cases, in times of high volatility this importance decreases for some "receivers", whereas for others it increases.

An intuition for these results might be found in the interconnection and commonalities of each country and the US. Specifically, factors such as trade, policy coordination or institutional similarities might be one reason for the industrial countries Australia, Canada, Japan and Korea to predominantly identify information from stock market fluctuations in the US. The US signal bears highly relevant and well-understood information that outweighs the uncertainty, and, is priced instantaneously. By contrast, the reduction of spillover intensity to the emerging economies Argentina, Brazil, Mexico and the Philippines in times of rising US volatility may be explained in the light of dissimilarities, for instance, in the institutional, legal and regulatory framework and relative political and economic instability. The information content in US price changes becomes less visible during turbulent times, which are perceived as propagating uncertainty instead.

4.4.4 Crisis, Correlation and Coefficients

During turbulent times, such as the ongoing global financial crisis, stock market co-movement is commonly perceived to be more pronounced. Indeed, splitting the present sample in a pre- and post-Lehman period with break date 9/15/2008 reveals a substantial increase in the empirical return correlation between each country and the US. Yet, at the same time, our previous results showed *decreasing* spillover intensity in some markets (Argentina, Brazil, Mexico and the Philippines). Even though we already specified a time varying coefficient model, these findings suggest that the volatility effect on the transmission strength might exhibit a structural break. So far, our approach implicitly assumed that either the information or the uncertainty hypotheses predominates over the whole sample period. Therefore, we pursue this issue further with emphasis on a pre-crisis and a crisis sample.

It is also well known, however, that a rise in correlation between two variables might very well simply be triggered by an increased variance of the explanatory variable. Forbes and Rigobon (2002), for instance, document this crucial role of volatility changes that can result in biased estimates of correlation coefficients. For the present data we evaluated this effect in a small simulation study. Denoting US returns by x_t and those of the other country by y_t, we simulated $y_t = \beta x_t + \epsilon_t$ for the pre- and post-Lehman period with parameters according to our empirical estimates from the above models. Thereby, the following rule of thumb was used: We set β to the average spillover intensity and drew ϵ_t and x_t from normal distributions with zero mean and $\text{Var}(\epsilon_t)$ and $\text{Var}(x_t)$ equal to the average ARCH variances - before and after 9/15/2008, respectively.

With this parametrization we were able to reproduce the sharp rise in return correlation during the crisis period. Thus, the increasing US volatility turned out to be the major driving force behind the rising correlations with Argentina, Brazil, Mexico and the Philippines. At the same time, this implies that the transition functions with stable parameters are compatible with the data. Despite the increase in return correlations, our approach is able to identify what we have termed the uncertainty effect, i.e., spillover strength decreases in volatility. The reason is that the variance changes, which affect the correlation coefficients, are explicitly taken into account in our model.

4.5 Conclusion

The present study motivated volatility-dependent simultaneous stock market interaction by discussing the fundamental character of volatility, which we argue is inherently ambivalent. Regarding the academic literature, volatility is used to proxy two different latent variables: information and uncertainty. We summarize the first view as the information hypothesis referring to studies where volatility is directly related to information flow intensity (see e.g. Ross 1989, Foster and Viswanathan 1993, 1995 or Kalev et al. 2004). The uncertainty hypothesis, on the other hand, has its source in large strands of literature where volatility is functioning as an uncertainty-proxy (see e.g Engle et al. 1987, Grier and Perry 2000, Kiyota and Urata 2004, Bekaert et al. 2009 or Li 2011).

We propose an econometric approach that consists of a simultaneous equations model with time-varying parameters. The time-variation of the spillover coefficient in one market is driven by the volatility of the other. In this setting it is the effect of volatility on the spillover strength that reflects whether the information hypothesis (positive effect) or the uncertainty hypothesis (negative effect) dominates.

Our main finding is that stock market interaction depends significantly on volatility in all countries under investigation. Evidence for the information hypothesis is found for the industrial countries (Australia, Canada, Japan and Korea), whereas the data of developing countries (Argentina, Brazil, Mexico and the Philippines) support the uncertainty hypothesis.

This paper reveals that foreign volatility plays a crucial role in the interaction of stock markets. Thereby, the signal of volatility differs substantially across countries. We show that, apart from the well-known capability of conditional variances to capture volatility clusters and ensuring efficient estimation, they constitute a useful tool for further purposes. Namely, conditional variances also help identify simultaneous effects and, especially, describe the time-varying nature of these effects in financial applications.

4.A Constant, Linear and Non-Linear Spillover

Table 4.2: Constant, Linear, and Non-Linear Spillover

	constant coefficients	*linearly time-varying coefficients*		*non-linearly time-varying coefficients*			
spillover on	a_{ij}	a_{ij}	b_{ij}	a_{ij}	$b_{ij,1}$	$b_{ij,2}$	$b_{ij,3}$
US	−0.24	−0.05	−0.50	−0.01	−0.52	−0.01	0.06
	[0.07]	[0.06]	[0.09]	[0.07]	[0.13]	[0.00]	[0.11]
Canada	−0.27	−0.20	−0.27	0.04	−1.33	1.64	−0.72
	[0.03]	[0.04]	[0.06]	[0.13]	[0.72]	[1.24]	[0.65]
US	–	–	–	–	–	–	–
Australia	−0.39	−0.35	−0.08	−0.33	−0.05	−0.23	0.14
	[0.01]	[0.02]	[0.04]	[0.10]	[0.44]	[0.58]	[0.21]
US	–	–	–	–	–	–	–
Japan	−0.42	−0.35	−0.14	−0.17	−0.77	0.53	−0.10
	[0.01]	[0.03]	[0.05]	[0.14]	[0.58]	[0.73]	[0.26]
US	–	–	–	–	–	–	–
Korea	−0.35	−0.29	−0.09	−0.06	−0.72	0.18	0.12
	[0.02]	[0.04]	[0.06]	[0.16]	[0.72]	[0.92]	[0.34]
US	0.02	0.04	−0.01	0.17	−0.11	0.02	−0.00
	[0.00]	[0.01]	[0.00]	[0.03]	[0.03]	[0.01]	[0.00]
Argentina	−0.69	−0.97	0.36	−1.84	2.93	−2.42	0.68
	[0.02]	[0.05]	[0.07]	[0.24]	[0.89]	[1.04]	[0.35]
US	0.02	0.04	−0.01	−0.02	0.10	−0.06	0.01
	[0.00]	[0.01]	[0.01]	[0.02]	[0.01]	[0.00]	[0.00]
Brazil	−0.80	−0.97	0.22	−1.58	2.60	−2.57	0.78
	[0.02]	[0.06]	[0.08]	[0.24]	[0.92]	[1.08]	[0.38]
US	−0.01	0.04	−0.05	0.39	−1.00	0.80	−0.20
	[0.01]	[0.03]	[0.02]	[0.11]	[0.31]	[0.30]	[0.09]
Mexico	−0.68	−0.77	0.11	−0.90	0.47	−0.33	0.09
	[0.03]	[0.05]	[0.04]	[0.15]	[0.55]	[0.63]	[0.21]
US	–	–	–	–	–	–	–
Philippines	−0.42	−0.46	0.07	−0.35	−0.65	1.26	−0.54
	[0.02]	[0.03]	[0.04]	[0.14]	[0.55]	[0.65]	[0.23]

Notes: This table shows results from a constant coefficient model, from a linear transition function and from a non-linear transition function. Standard errors are given in brackets. The estimated coefficients refer to the transition functions $A_{ijt} = a_{ij}$, $A_{ijt} = a_{ij} + b_{ij}h_{jt}$ and $A_{ijt} = a_{ij} + b_{ij,1}h_{jt} + b_{ij,2}h_{jt}^2 + b_{ij,3}h_{jt}^3$ of the simultaneous model:

$$\begin{pmatrix} 1 & A_{12t} \\ A_{21t} & 1 \end{pmatrix} y_t = \varepsilon_t \ .$$

Bibliography

Adrian, T., H. Wu (2009): The Term Structure of Inflation Expectations. *Federal Reserve Bank of New York Staff Report*, no. 362.

Andersen, T.G. (1996): Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. *Journal of Finance*, 51, 169–204.

Andersen, T.G., T. Bollerslev (1998): Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies. *Journal of Finance*, 53, 219–265.

Ang, A., M. Piazzesi (2003): A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics*, 50, 745–787.

Aydemir, A.C. (2008): Risk Sharing and Counter-Cyclical Variation in Market Correlations. *Journal of Economic Dynamics & Control*, 32, 3084–3112.

Baillie, R.T., T. Bollerslev, H.O. Mikkelsen (1996): Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 74, 3–30.

Balfoussia, H., M.R. Wickens (2007): Macroeconomic Sources of Risk in the Term Structure. *Journal of Money, Credit and Banking*, 39, 205–236.

Bali, T.G., R.F. Engle (2010): The Intertemporal Capital Asset Pricing Model with Dynamic Conditional Correlations. *Journal of Monetary Economics*, 57, 377–390.

Banerjee, A., J. Dolado, R. Mestre (1998): Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework. *Journal of Time Series Analysis*, 19, 267–283.

Beechey, M.J., B.K. Johannsen, A. Levin (2011): Are Long-Run Inflation Expectations Anchored More Firmly in the Euro Area than in the United States? *American Economic Journal: Macroeconomics*, 3, 104–129.

Bekaert, G., E. Engstrom, Y. Xing (2009): Risk, Uncertainty, and Asset Prices. *Journal of Financial Economics*, 91, 59–82.

Bernanke, B.S., V.R. Reinhart, B.P. Sack (2004): Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment. *Brookings Papers on Economic Activity*, 2, 1–100.

BoE (2010): *Inflation Report May 2010*, United Kingdom.

Bollerslev, T. (1986): Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., R.F. Engle, J.M. Wooldridge (1988): A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, 96, 116–131.

Bollerslev, T., J.M. Wooldridge (1992): Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. *Econometric Reviews*, 11, 143–172.

Bremnes, H., G. Oystein, S. Frode (2001): Linkages among Interest Rates in the United States, Germany and Norway. *Scandinavian Journal of Economics*, 103, 127–145.

Busch, T. (2005): A Robust LR Test for the GARCH Model. *Economics Letters*, 88, 358–364.

Campbell, J.Y., R.J. Shiller (1987): Cointegration and Tests of Present Value Models. *Journal of Political Economy*, 95, 1062–1088.

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

Campbell, J.Y., J.H. Cochrane (1999): By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107, 205–251.

Caporale, B., T. Caporale (2008): Political Risk and the Expectations Hypothesis. *Economics Letters*, 100, 178–180.

Carstensen, K. (2003): Nonstationary Term Premia and Cointegration of the Term Structure. *Economics Letters*, 80, 409–413.

Charles, A., O. Darné (2008): A Note on Unit Root Tests and GARCH Errors: A Simulation Experiment. *Communications in Statistics - Simulation and Computation*, 37, 314–319.

Chowdhury, A.R. (1993): Does Exchange Rate Volatility Depress Trade Flows? Evidence from Error-Correction Models. *Review of Economics and Statistics*, 75, 700–706.

Christensen, J.H.E., J.A. Lopez, G.D. Rudebusch (2010): Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields. *Journal of Money, Credit and Banking*, 42, 143-178.

Christiansen, C. (2000): Macroeconomic Announcement Effects on the Covariance Structure of Government Bond Returns. *Journal of Empirical Finance*, 7, 479–507.

Clark, P.K. (1973): A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41, 135–155.

Cochrane, J. (2001): *Asset Pricing*. Princeton University Press, Princeton, New Jersey.

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

Connolly, R.A., Z.N. Güner, K.N. Hightower (2007): Evidence on the Extent and Potential Sources of Long Memory in U.S. Treasury Security Returns and Yields. *Journal of Money, Credit and Banking*, 39, 689–702.

Cox, J.C., J.E. Ingersoll, S.A. Ross (1985): A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385–407.

Cukierman, A., A.H. Meltzer (1986): A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information. *Econometrica*, 54, 1099–1128.

Cushman, D.O. (1985): Real Exchange Rate Risk, Expectations, and the Level of Direct Investment. *Review of Economics and Statistics*, 67, 297–308.

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

Demos, A. (2002): Moments and Dynamic Structure of a Time-Varying Parameter Stochastic Volatility in Mean Model. *Econometrics Journal*, 5, 345–357.

Dickey, D.A., W.A. Fuller (1979): Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, 427–431.

van Dijk, D., P.H. Franses, H.P. Boswijk (2007): Absorption of Shocks in Non-Linear Autoregressive Models. *Computational Statistics & Data Analysis*, 51, 4206–4226.

ECB (2011): *Monthly Bulletin April*, Frankfurt, Germany.

Elliott, G., T.J. Rothenberg, J.H. Stock (1982): Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64, 813–836.

Engle, R.F. (1982): Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987–1007.

Engle, R.F., Bollerslev, T. (1986): Modelling the Persistence of Conditional Variances. *Econometric Reviews*, 5, 1–50.

Engle, R.F., D.M. Lilien, R.P. Robins (1987): Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica*, 55, 391–407.

Engsted, T., C. Tanggaard (1994): Cointegration and the US Term Structure. *Journal of Banking & Finance*, 18, 167–181.

Epps, T.W., M.L. Epps (1976): The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-Of-Distributions Hypothesis. *Econometrica*, 44, 305–321.

Ericsson, N.R., J.G. MacKinnon (2002): Distributions of Error Correction Tests for Cointegration. *Econometrics Journal*, 5, 285–318.

Evans, M.D.D., K.K. Lewis (1994): Do Stationary Risk Premia Explain It All?: Evidence from the Term Structure. *Journal of Monetary Economics*, 33, 285–318.

Fleming, J., C. Kirby, B. Ostdiek (1998): Information and Volatility Linkages in the Stock, Bond, and Money Markets. *Journal of Financial Economics*, 49, 111–137.

Forbes, K.J., R. Rigobon (2002): No Contagion, Only Interdependence: Measuring Stock Market Comovements. *Journal of Finance*, 57, 2223–2261.

Foster, F.D., S. Viswanathan (1993): Variations in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models. *Journal of Finance*, 48, 187–211.

Foster, F.D., S. Viswanathan (1995): Can Speculative Trading Explain the Volume-Volatility Relation? *Journal of Business & Economic Statistics*, 13, 379–396.

Fountas, S., M. Karanasos, A. Mendoza (2004): Output Variability and Economic Growth: the Japanese Case. *Bulletin of Economic Research*, 56, 353–363.

Fountas, S., M. Karanasos (2007): Inflation, Output Growth, and Nominal and Real Uncertainty: Empirical Evidence for the G7. *Journal of International Money and Finance*, 26, 229–250.

French, K.R., R. Roll (1986): Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics*, 17, 5–26.

French, K.R., G.W. Schwert, R.F. Stambaugh (1987): Expected Stock Returns and Volatility. *Journal of Financial Economics*, 19, 3–29.

Friedman, M. (1977): Nobel Lecture: Inflation and Unemployment. *Journal of Political Economy*, 85, 451–472.

Gagnon, L., G.A. Karolyi (2009): Information, Trading Volume, and International Stock Return Comovements: Evidence from Cross-Listed Stocks. *Journal of Financial and Quantitative Analysis*, 44, 953–986.

Gefang, D., G. Koop, S.M. Potter (2012): The Dynamics of UK and US Inflation Expectations. *Computational Statistics and Data Analysis*, 56, 3120–3133.

Ghysels, E., P. Santa-Clara, R. Valkanov (2005): There is a Risk-Return Trade-Off After All. *Journal of Financial Economics*, 76, 509–548.

de Goeij, P., W. Marquering (2006): Macroeconomic Announcements and Asymmetric Volatility in Bond Returns. *Journal of Banking & Finance*, 30, 2659–2680.

Gordon, S., P. St-Amour (2004): Asset Returns and State-Dependent Risk Preferences. *Journal of Business & Economic Statistics*, 22, 241–252.

Greenwood, R., D. Vayanos (2010): Price Pressure in the Government Bond Market. *American Economic Review: Papers and Proceedings*, 100, 585–590.

Greenwood, R., D. Vayanos (2012): Bond Supply and Excess Bond Returns. *NBER Working Paper Series No. 13806.*

Grier, K.B., M.J. Perry (2000): The Effects of Real and Nominal Uncertainty on Inflation and Output Growth: Some GARCH-M Evidence. *Journal of Applied Econometrics*, 15, 45–58.

Gürkaynak, R.S., B. Sack, J.H. Wright (2007): The U.S. Treasury Yield Curve: 1961 to the Present. *Journal of Monetary Economics*, 54, 2291–2304.

Gürkaynak, R.S., B. Sack, J.H. Wright (2010a): The TIPS Yield Curve and Inflation Compensation. *American Economic Journal: Macroeconomics*, 2, 70–92.

Gürkaynak, R.S., E. Swanson, A. Levin (2010b): Does Inflation Targeting Anchor Long-run Inflation Expectations? Evidence from the U.S., UK, and Sweden. *Journal of the European Economic Association*, 8, 1208–1242.

Gürkaynak, R.S., J.H. Wright (2012): Macroeconomics and the Term Structure. *Journal of Economic Literature*, 50, 331–367.

Guibaud, S., Y. Nosbusch, D. Vayanos (2013): Bond Market Clienteles, the Yield Curve, and the Optimal Maturity Structure of Government Debt. *Review of Financial Studies*, Online ISSN 0893-9454.

Hall, A.D., H.M. Anderson, C.W.J. Granger (1992): A Cointegration Analysis of Treasury Bill Yields. *The Review of Economics and Statistics*, 74, 116–126.

Hansen, P.R. (2003): Structural Changes in the Cointegrated Vector Autoregressive Model. *Journal of Econometrics*, 114, 261–295.

Harris, L. (1987): Transaction Data Tests of the Mixture of Distributions Hypothesis. *Journal of Financial and Quantitative Analysis*, 22, 127–141.

Hassler, U., J. Wolters (1995): Long Memory in Inflation Rates: International Evidence. *Journal of Business & Economic Statistics*, 13, 37–45.

Hecq, A. (1996): IGARCH Effect on Autoregressive Lag Length Selection and Causality Tests. *Applied Economics Letters*, 3, 317–23.

Hess, A.C., A. Kamara (2005): Conditional Time-Varying Interest Rate Risk Premium: Evidence from the Treasury Bill Futures Market. *Journal of Money, Credit and Banking*, 37, 679–698.

Ibbotson and Associates (1994): *Stocks, Bonds, Bills, and Inflation Yearbook.* Ibbotson Associates, Chicago.

Jochmann, M., G. Koop, S.M. Potter (2010): Modeling the Dynamics of Inflation Compensation. *Journal of Empirical Finance*, 17, 157–167.

Johnson, P.A. (1994): On the Number of Common Unit Roots in the Term Structure of Interest Rates. *Applied Economics*, 26, 815–20.

Jones, C.M., O. Lamont, R.L. Lumsdaine (1998): Macroeconomic News and Bond Market Volatility. *Journal of Financial Economics*, 47, 315–337.

Kalev, P.S., W.-M. Liu, P.K. Pham, E. Jarnecic (2004): Public Information Arrival and Volatility of Intraday Stock Returns. *Journal of Banking & Finance*, 28, 1441–1467.

Kapetanios, G., Y. Shin, A. Snell (2003): Testing for a Unit Root in the Non-Linear STAR Framework. *Journal of Econometrics*, 112, 359–379.

Kilian, L., M.P. Taylor (2003): Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates? *Journal of International Economics*, 60, 85–107.

Kim, O., R. Verrecchia (1991): Trading Volume and Price Reaction to Public Announcements. *Journal of Accounting Research*, 29, 302–321.

Kim, K., P. Schmidt (1993): Unit Root Tests with Conditional Heteroskedasticity. *Journal of Econometrics*, 59, 287–300.

Kim, W., O. Linton (2011): Estimation of a Semiparametric IGARCH(1,1) Model. *Econometric Theory*, 27, 639–661.

King, M.A., S. Wadhwani (1990): Transmission of Volatility between Stock Markets. *Review of Financial Studies*, 3, 5–33.

Kiyota, K., S. Urata (2004): Exchange Rate, Exchange Rate Volatility and Foreign Direct Investment. *The World Economy*, 27, 1501–1536.

Koop, G., H. Pesaran, S.M. Potter (1996): Impulse Response Analysis in Non-Linear Multivariate Models. *Journal of Econometrics*, 74, 199–147.

Kourogenis, N., N. Pittis (2008): Testing for a Unit Root Under Errors with Just Barely Infinite Variance. *Journal of Time Series Analysis*, 29, 1066–1087.

Kruse, R., P. Sibbertsen (2011): Long Memory and Changing Persistence. *Economics Letter*, 114, 268–272.

Kugler, P. (1988): An Empirical Note on the Term Structure and Interest Rate Stabilization Policies. *The Quarterly Journal of Economics*, 103, 789–792.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, Y. Shin (1992): Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure are We That Economic Time Series Have a Unit Root? *Journal of Econometrics*, 54, 159–178.

Kyle, A. (1985): Continuous Auctions and Insider Trading. *Econometrica*, 53, 1315–1335.

Lamoureux, C.G., W.D. Lastrapes (1990a): Persistence in Variance, Structural Change, and the GARCH Model. *Journal of Business & Economic Statistics*, 8, 225–234.

Lamoureux, C.G., W.D. Lastrapes (1990b): Heteroskedasticity in Stock Return Data: Volume Versus GARCH Effects. *Journal of Finance*, 45, 221–229.

Levin, A.T., F.N. Natalucci, J.M. Piger (2004): Explicit Inflation Objectives and Macroeconomic Outcomes. *Working Paper Series 383*, European Central Bank.

Li, J. (2011): Volatility Components, Leverage Effects, and the Return-Volatility Relations. *Journal of Banking & Finance*, 35, 1530–1540.

Ling, S., M. McAleer (2003): Estimation and Testing for Unit Root Processes with GARCH (1, 1) Errors: Theory and Monte Carlo Evidence. *Econometric Reviews*, 22, 179–202.

Lintner, J. (1965): The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47, 13–37.

Lucas, R.E. (1978): Asset Prices in an Exchange Economy. *Econometrica*, 46, 1429–1445.

Lumsdaine, R.L. (1995): Finite-Sample Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models: A Monte Carlo Investigation. *Journal of Business & Economic Statistics*, 13, 1–10.

Lundblad, C. (2007): The Risk Return Tradeoff in the Long Run: 1836-2003. *Journal of Financial Economics*, 85, 123–150.

Luukkonen, R., P. Saikkonen, T. Teräsvirta (1988): Testing Linearity Against Smooth Transition Autoregressive Models. *Biometrika*, 75, 491–499.

Mankiw, N.G., J.A. Miron (1986): The Changing Behavior of the Term Structure of Interest Rates. *The Quarterly Journal of Economics*, 101, 211–228.

Mele, A. (2007): Asymmetric Stock Market Volatility and the Cyclical Behavior of Expected Returns. *Journal of Financial Economics*, 86, 446–478.

Meller, B., D. Nautz (2012): Inflation Persistence in the Euro Area Before and After the European Monetary Union. *Economic Modelling*, 29, 1170–1176.

Modigliani, F., R. Sutch (1966): Innovations in Interest Rate Policy. *American Economic Review*, 56, 178–197.

Nelson, D.B. (1990): Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, 6, 318–334.

Nelson, D.B. (1991): Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59, 347–370.

Nelson, R.C., F.A. Siegel (1987): Parsimonious Modeling of Yield Curves. *Journal of Business*, 60, 473–489.

Newey, W.K., K.D. West (1987): A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703–708.

Nobay, B., I. Paya, D.A. Peel (2010): Inflation Dynamics in the U.S.: Global but Not Local Mean Reversion. *Journal of Money, Credit and Banking*, 42, 135-150.

Ohanissian, A., J.R. Russell, R.S. Tsay (2008): True or Spurious Long Memory? A New Test. *Journal of Business & Economic Statistics*, 26, 161–175.

Pagan, A.R., A.D. Hall, V. Martin (1996): Modelling the Term Structure. *Handbook of Statistics*, 91–118.

Reinhart, V.R., B.P. Sack (2000): The Economic Consequences of Disappearing Government Debt. *Brookings Papers on Economic Activity*, 2, 163–220.

Rigobon, R. (2003): Identification Through Heteroskedasticity. *Review of Economics and Statistics*, 85, 777–792.

Rosenberg, J.V., R.F. Engle (2002): Empirical Pricing Kernels. *Journal of Financial Economics*, 64, 341–372.

Ross, S.A. (1989): Information and Volatility: The No-Arbitrage Martingale Approach to Timing and Resolution Irrelevancy. *Journal of Finance*, 44, 1–17.

Rubinstein, M. (1976): The Valuation of Uncertain Income Streams and the Pricing of Options. *The Bell Journal of Economics*, 7, 407–425.

Saikkonen, P., R. Luukkonen (1988): Lagrange Multiplier Tests for Testing Non-Linearities in Time Series Models. *Scandinavian Journal of Statistics*, 15, 55–68.

Sentana, E., G. Fiorentini (2001): Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models. *Journal of Econometrics*, 102, 143–164.

Seo, B. (1999): Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity. *Journal of Econometrics*, 91, 113–144.

Seo, B. (2007): Asymptotic Distribution of the Cointegrating Vector Estimator in Error Correction Models with Conditional Heteroskedasticity. *Journal of Econometrics*, 137, 68–111.

Sharpe, W.F. (1964): Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19, 425–442.

Shea, G.S. (1992): Benchmarking the Expectations Hypothesis of the Interest-Rate Term Structure: An Analysis of Cointegration Vectors. *Journal of Business & Economic Statistics*, 10, 347–366.

Shiller, R.J. (1979): The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure. *Journal of Political Economy*, 87, 1190–1219.

Shiller, R.J. (1990): The Term Structure of Interest Rates. In B.M. Friedman, F.H. Hahn, editors, *Handbook of Monetary Economics*, 627–722. North-Holland, Amsterdam.

Silverman, B.W. (1986): *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London, UK.

Skalin, J. (1998): Testing Linearity Against Smooth Transition Autoregression Using a Parametric Bootstrap. *SSE/EFI Working Paper in Economics and Finance 276, Stockholm School of Economics.*

Smith, P., M.R. Wickens (2002): Asset Pricing with Observable Stochastic Discount Factors. *Journal of Economic Surveys*, 16, 397–446.

Stock, J.H. (1987): Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica*, 55, 1035–1056.

Storn, R.M., K.V. Price (1997): Differential Evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11, 341–359.

Strohsal, T. (2009): Persistence in US Interest Rate Spreads: An Empirical Analysis. *Diploma Thesis*, Freie Universität Berlin, Berlin.

Strohsal, T., E. Weber (2013): Mean-Variance Cointegration and the Expectations Hypothesis. *Quantitative Finance*, forthcoming, doi:10.1080/14697688.2013.814974.

Svensson, L.E.O. (1994): Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. *National Bureau of Economic Research Working Paper 4871.*

Tauchen, G.E., M. Pitts (1983): The Price Variability-Volume Relationship on Speculative Markets. *Econometrica*, 51, 485–505.

Teräsvirta, T. (1994), Specifcation, Estimation and Evaluation of Smooth Transition Autoregressive Models. *Journal of the American Statistical Association*, 89, 208–218.

Tzavalis, E., M.R. Wickens (1995): The Persistence in Volatility of the US Term Premium 1970-1986. *Economics Letters*, 49, 381–389.

Vasicek, O. (1977): An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5, 177–188.

Vayanos, D., J.-L. Vila (2009): A Preferred-Habitat Model of the Term Structure of Interest Rates. *Discussion paper 641. Financial Markets Group, London School of Economics and Political Science*, London, UK.

Weber, E. (2010): Structural Conditional Correlation. *Journal of Financial Econometrics*, 8, 392–407.

Wolters, J. (1995): On the Term Structure of Interest Rates – Empirical Results for Germany. *Statistical Papers*, 36, 193–214.

Wolters, J. (1998): Cointegration and German Bond Yields. *Applied Economics Letters*, 5, 497–502.

Zhang, H. (1993): Treasury Yield Curves and Cointegration. *Applied Economics*, 25, 361–67.

Zivot, E., W.K. Andrews (1992): Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, 20, 25–44.

Ehrenwörtliche Erklärung

Ich habe die vorgelegte Dissertation selbst verfasst und dabei nur die von mir angegebenen Quellen und Hilfsmittel benutzt. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, sowie alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Berlin, den 29. August 2013