Decision-making, uncertainty and the predictability of financial markets

Essays on interest rates, crude oil prices and exchange rates

DISSERTATION
ZUR ERLANGUNG DES DOKTORGRADS
DER WIRTSCHAFTSWISSENSCHAFTLICHEN FAKULTÄT
DER GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

VORGELEGT VON

FREDERIK KUNZE
geboren in Neustadt am Rübenberge

GÖTTINGEN, 2018

Erstgutachter:	Prof. Dr. Kilian Bizer
Zweitgutachter:	Prof. Dr. Markus Spiwoks
Weiteres Mitglied der Prüfungskommission:	Prof. Dr. Jan Muntermann

Non-technical Summary

Decision-makers are confronted with decisions under uncertainty. Financial uncertainty may adversely affect growth. Theoretically, forecasts may potentially reduce uncertainty and create economic value. Focusing on survey predictions, this cumulative dissertation addresses the economic relevance of interest rate, crude oil and exchange rate forecasts for policy as well as managerial decision-makers and financial market participants, respectively. The first research objective of the presented studies is to compile novel evidence on the accuracy, rationality and usefulness of financial market forecasts delivered by professional analysts. Despite the comprehensible critique regarding their foresight qualities in efficient markets, financial forecasts are to be understood as integral elements for decision-makers of various kinds and hence may even be classified as indispensable. This makes the appraisal of forecasts and the corresponding price-buildings processes even more essential for decision-makers. Following this line of thought and focusing again on interest rates, crude oil prices and exchange rates, the second research objective of this thesis is devoted to the analysis of changing market environments and the resulting effects on expectation formations.

Nicht-technische Zusammenfassung

Es liegt in der Verantwortung von ökonomischen Entscheidungsträgern, Entscheidungen unter Unsicherheit zu treffen. Finanzmarktunsicherheit kann ökonomische Aktivität negativ beeinflussen. In diesem Sinn können Prognosen theoretisch einen ökonomischen Mehrwert liefern. Diese kumulative Dissertation beschäftigt sich mit der ökonomischen Relevanz von Vorhersagen für Zinsen, Rohölpreise und Wechselkurse für Entscheidungsträger aus Politik und Wirtschaft sowie für Finanzmarktteilnehmer. Der Fokus dieser Arbeit liegt dabei auf umfragebasierten Survey-Prognosen. Der erste Forschungsbeitrag dieser Arbeit bezieht sich auf neue empirische Erkenntnisse mit Blick auf die Genauigkeit, Rationalität und Verwendbarkeit von von Finanzmarktexperten zur Verfügung gestellten Prognosen für Finanzmarktzeitreihen. Trotz der nachvollziehbaren Kritik mit Blick auf die Prognosequalität in effizienten Märkten sind Finanzmarktprognosen essentielle Entscheidungskriterien für die handelnden Akteure. Insofern sind Finanzmarktprognosen unverzichtbar, was aus Sicht der Entscheidungsträger eine intensive Begutachtung der Prognosen sowie der den Finanzmarktzeitreihen zu Grunde liegenden Preisbildungsmechanismen erforderlich macht. Auf diesem Gedankengang aufbauend leitet sich der zweite Forschungsbeitrag dieser Arbeit ab, der sich mit den sich verändernden Rahmenbedingungen in ausgewählten Zins-, Rohöl- und Devisenmärkten beschäftigt.

Sabrina, Florentine, Henriette
and my parents

Contents

Chapter 1

Introduction and Summaries

"Extrapolations into the future, whether empirically founded projections or hunches that are revealed only by their consequences in action, are being made and will continue to be made. We live in the present and cannot avoid the future; the decisions we make today will affect tomorrow. Indeed, many of them must look toward a longer-range future. Such decisions, whether made against the background of articulated forecasts or out of a subconscious but often quite strong feeling about the climate of life to be expected, imply projection in the sense of some view of the future. The choice is not between making and not making an extrapolation into the future; it is between making the projection in overt and sometimes quantitative terms, and proceeding by feel and by faith. Even inaction implies some picture of the future."

Simon Kuznets (1954)[1]

Introduction

Both corporate and political decision-makers as well as financial market participants and private households are confronted with decisions under uncertainty.[2] For decades, economic scholars have been calling attention to uncertainty related impediments for real economic activity.[3] Following this line of thought, it has been argued that rising financial uncertainty may depress inter alia investement activity and the supply of credit and hence adversely affect growth.[4] Consequently, it has been intensively discussed in the academic literature that economic and financial market forecasts may potentially reduce uncertainty (Holden et al., 1990), lead to *"good decisions"* (see Diebold and Lopez, 1996, p. 241) and hence create economic value.[5] Owing to this, the usefulness of forecasting methods for economic indicators as well as financial market variables and their accuracy has been empirically assessed and critically discussed for several decades (See Cox, 1930; Makridakis et al., 1979; Hendry and Clements, 2003; Rossi and Sekhposyan, 2016, to name but a few).

Unsurprisingly, the forecast evaluation literature has rapidly grown in the last three decades (see, e.g, Diebold and Lopez, 1996; West, 2006; Franses et al., 2014, who provide comprehensive overviews of empirical studies and methodologies). In this context, it is a valid and well-known assertion that the forecast profession

1 See Kuznets (1954), p. 36. Also partly quoted in Roos (1955).
2 See, for example, Lowe (1970) and Belongia (1987).
3 See, especially, Keynes (1936) and Friedman (1977) as well as Creal and Wu (2017) and Moore (2017) for more recent investigations in this regard.
4 See, for example, Caldara et al. (2016).
5 See, for example, Howard (1954), Wheelwright and Clarke (1976) as well as Elliott and Timmermann (2008).

always has had to deal with the critique that predictions are useless[6] or inaccurate at least. In recent times, not only the limitations regarding the accuracy of forecasts but also forecasters' behavioral patterns have been discussed intensively.[7]

Furthermore, the bounded utility of forecasts receives much more attention in times of financial and economic crisis.[8] In this regard, reasonable complaints are that forecasts seem to be least accurate when they are needed most[9] and that there exist significant limitations of prediction methods in anticipating crisis events or recessions (see, for example, Fintzen and Stekler, 1999; Goodwin and Wright, 2010). Rising forecast errors during regime shifts as well as in times of high uncertainty and recession events do often correspond with structural breaks in long-term relationships (Stock and Watson, 1996; Pesaran and Timmermann, 2004; Clements and Hendry, 2006), which are important elements of almost all forecasting approaches.[10] Moreover, survey forecasts for economic indicators and financial variables – serving as a rare proxy for market expectations – seem to regularly violate the important assumption of rationality. Hence, it is realistic and advisable to ask whether there is any use of forecasts at all.

Having said that, decision-making – which by definition describes a structural process to make plans for coming periods – is difficult to imagine without assumptions about the future (Kuznets, 1954). Phrased somewhat differently: forward-looking planning and strategies in fact need forecasting (Roos, 1955; Schultz, 1984; Makridakis, 1986) or, following Firth (1975), *"Planning involves making decisions which will have their effect in the future and so an estimate of this future is required. This assessment of the future is termed forecasting and it is a vital ingredient in any planning process."* (see Firth, 1975, p. 97). In fact, forecasting may not be seen as an unsolicitous task for decision-makers. This is especially true for financial institutions. In order to be compliant with international accounting standards – i.e. *International Financial Reporting Standard 9* (IFRS 9) – banks are obliged to make use of economic forecasts and financial market predictions (i.e. interest rates) and employ macroeconomic-based credit loss models (Novotny-Farkas, 2016; Feschiyan and Andasarova, 2017; Skoglund, 2017).

Under the assumption of an implicit accordance that despite apparent deficiencies it is not possible to imagine decision-making without forecasting, it is even

6 Especially under the assumption of efficient financial markets (Mitchell and Pearce, 2007; Gubaydullina et al., 2011; Baghestani et al., 2015). Timmermann and Granger (2004) argued: *"The efficient market hypothesis (EMH) is a backbreaker for forecasters"* (see Timmermann and Granger, 2004, p. 15).

7 For example, Meub et al. (2013) experimentally confirmed that anchoring leads to biased forecasts. Also focusing on the behavioral biases of forecasters, Proeger and Meub (2014) examined the influence of forecasters' overconfidence and concluded that this is more of a social rather than only an individual bias. Focusing on survey predictions for financial markets, Fujiwara et al. (2013) emphasized that professional forecasters are influenced by past predictions and hence act behaviorally.

8 For example, Dua (1988) mentioned a decline in accuracy of interest rate forecasts in periods of high volatility.

9 Crisis-related effects on forecast accuracy have been inter alia discussed by Kunze and Gruppe (2014), Kunze (2014) as well as Kunze et al. (2015).

10 It does not necessarily have to be the case that crisis events lead to structural changes in long-term relationships. In his famous *Lucas Critique*, Lucas Jr (1976) articulated the concern that with their decisions policy makers themselves *"will systematically alter the structure of econometric models"* (see Lucas Jr, 1976, p. 41). This endogeneity problem is especially relevant for central banks, which is outlined in further detail below.

more important to have a clear understanding of the methodological issues of financial forecasting and the corresponding price-building processes.[11] Furthermore, one important characteristic aftermath of financial crises is the increased attention of decision-makers as well as researchers and forecasters regarding the interdependencies of financial market turmoil, macroeconomic uncertainty and economic growth (Bekaert et al., 2013; Creal and Wu, 2017). This has also been true for the global financial crisis as well as the subsequent euro crisis,[12] because the causes and consequences of these crises demonstrated quite impressively that the pro-cyclical features and the fragility of financial markets – with rising financial imbalances and following corrections of asset prices – also hold strong relevance for real economic activity[13] and hence for economic and financial forecasts as well as their recipients.

Motivation and objectives of this thesis

Added together, these preliminary remarks inevitably lead to the conclusion that given the postulated relationship between economic activity and financial market uncertainty financial market forecasts are indeed indispensable ingredients of almost any decision-making process in business and economics, but they have to be appraised in the light of a constantly-changing environment. The adjustments of foreign exchange rate regimes in many Asian economies as a response to the Asian crisis in 1997/98 may be seen as one well-fitting example in this context (Rajan, 2002; Hernandez and Montiel, 2003). However, despite being regularly associated with crisis events, structural changes might also be a consequence of significantly altered institutional circumstances and hence policy driven. In this context, the introduction of the euro and the subsequent convergence of government bond yields of the European Monetary Union (EMU) member countries, the opening up of China's financial markets and the corresponding internationalization of the Chinese currency as well as the recently observed changes to the balance of power regarding global oil markets – also referred to as the *New Oil Order*[14] – are relevant "real life" examples, that will also be touched in this thesis.

Focusing on survey predictions, this thesis' studies address the economic relevance of interest rate, crude oil price and exchange rate forecasts for policy as well as managerial decision-makers and financial market participants, respectively. The first research objective of the presented studies is to compile novel evidence on the accuracy, rationality and usefulness of financial market forecasts delivered by professional analysts. Despite the comprehensible critique regarding to their foresight

[11] In this regard, the necessity to recognize the characteristics of exchange rate regimes to predict exchange rate movements is only one prominent example (see, for example, von Spreckelsen et al., 2014).

[12] See, for example, Bloom (2009), Beckmann and Czudaj (2017a), Creal and Wu (2017), and Hartmann et al. (2017)

[13] See, for example, Gramlich and Oet (2011), Borio (2014) and Magkonis and Tsopanakis (2016) as well as Cardarelli et al. (2011) who delivered a comprehensive literature overview regarding the impact of financial cycles on the real economy. The authors inter alia discussed the effects of eroding values of collateral and the preparedness of the financial system to grant funds for business activities as well as the consequences of eroding bank capital.

[14] In 2015, this expression was used by researchers at the investment bank Goldman Sachs (see Damie et al., 2015; Khan, 2017) and has to be seen closely associated with the shale oil boom in the US (See also Morecroft, 2017, who investigated the so-called *Saudi America* hypothesis.)

qualities in efficient markets[15] financial forecasts are indeed to be understood as integral elements for decision-makers of various kinds and hence may even be classified as indispensable. In this context, Granger and Pesaran (2000) and Timmermann and Granger (2004) annotated the need to link decision-making and forecast accuracy. Furthermore, the appraisal of the usefulness of these forecasts is always a matter of the specific task of the decision-makers and the circumstances in which they will be applied (Barron and Targett, 1988). Moreover – and against the background of constantly-changing market environments, periods of financial stress and recessions – survey forecasts may deliver useful insights for decision-makers that extent beyond pure measures of accuracy and rationality (see also Makridakis, 1996). In this regard, the potential relationship between uncertainty and financial forecasters' disagreement has received a substantial amount of attention lately and looking forward it could contribute to explain the relationship between financial market movements and the real economy. Reflecting this line of thought leads to the second research objective which this thesis will be devoted: the relevance of changing market environments and the resulting effects on expectation formation and price-building processes in financial markets.

Before summarizing the following chapters of this thesis, the remainder of this introductory chapter will discuss the relevance of financial forecasts for various decision-makers in an uncertain world and emphasize the relevance and applications for survey predictions for interest rates, crude oil prices and exchange rates.

Decision-makers and financial forecasts

In general, economic and financial forecasts should be interpreted as an integral part of any decision theoretical framework and hence they have no intrinsic value.[16] In practice, the potential contribution of forecasts largely depends on the specific objectives and tasks of decision-makers. It should not come as a surprise, that predictions for financial market variables are first and foremost linked to their aid regarding the implementation of successful trading strategies (see also Kim and Orphanides, 2012).[17] Discussing the relevance and potential benefits of forecasts for financial variables, Mills (2008) brought up a well-fitting example presented by Lo (1997), which seems to be worth revisiting at this point. The author compared an investment of one US dollar in January 1926 into US Treasury bills with a maturity of one month with the investment of the same amount in the S&P stock market index. Under the assumption of reinvestment, Lo (1997) stated that the bond investment would have grown to 12 US dollars by December 1994 whereas the stock investment would have reached 811 US dollars in the same period. Lo (1997) further developed his example by altering the reinvestment method: now, each month an investor with perfect foresight (i.e. with the knowledge concerning which asset would have the higher yield) could choose between the stock and the bond investment (see Lo, 1997; Mills, 2008). In this case, the final amount would be 1,251,684,443 US dollars.[18] Mills (2008) boiled Lo's finding down to the essence: *"Obviously, few, if any, investors have*

[15] See also Schwartz (1970), Fama (1970) as well as Belongia (1987).

[16] See McNees (1988), Winklhofer et al. (1996) as well as Elliott and Timmermann (2008).

[17] It is argued here that financial market forecasts are a specific sub-theme of economic forecasts.

[18] Using slightly different investment periods, Mills (2008) ended up with different results.. Moreover, it is important to note that this example ignores transaction costs (Mills, 2008) and does not account for survivorship bias. After all, the general implications of this numerical show piece remain unaltered.

perfect foresight, but Lo's point was that even modest ability to forecast financial asset returns would have been handsomely rewarded [...]" (Mills, 2008, p. 510).

In fact, not only speculative motives lead to a substantial demand for financial market forecasts. In this regard, Baghestani et al. (2015) emphasized the relevance of interest rate forecasts and the inference of market expectations for (monetary) policy decision-makers, whereas Duffee (2013) additionally highlighted the significance of interest rate forecasts as an input factor for economic forecasts. Alquist et al. (2013) annotated the impact of oil price forecasts on managerial decision-making and Giddy and Dufey (1975) remarked that foreign exchange rate forecasts have to be seen as *"significant inputs to decisions concerning practically every aspects of international business"* (see Giddy and Dufey, 1975, p. 1).

More generally, forecasts for interest rates, oil prices and exchange rates offer assistance for decision-makers who have to cope with uncertainty and decisions under risk. The remainder of this section provides a more detailed overview of potential recipients of forecasts.

Non-financial institutions – to begin with – are in fact exposed to financial risks. There exists ample empirical evidence in the literature regarding non-financial firms' interest rate (Bartram, 2002; Dhanani et al., 2007), exchange rate (Jorion, 1990; Bodnar and Gentry, 1993; Bartram et al., 2010) as well as commodity price exposure (Tsai, 2015; Shaeri et al., 2016). Concerning exchange rate risks, it is important to note that even without direct foreign business activities – in terms of their assets, liabilities or general operations – firms may be exposed to exchange rate risk.[19]

Consequently, effective risk management may increase firm value (Aretz et al., 2007). In order to mitigate these risks and increase firm value, hedging might be considered as the preferred strategy by corporate decision-makers. Froot et al. (1993) extensively discussed the benefits and rationales for hedging and inter alia annotated that hedging may be of benefit for corporations when external financing is more costly than internal funds. In fact, in the long-term corporate decision-makers could theoretically eliminate almost any exposure via hedging strategies. However, these hedging strategies are costly and hence corporate decision-makers must rely on forecasts to determine the optimal amount and form of hedging (Stockman, 1987). As a result, forecasting financial market variables is an integral part of corporate hedging strategies, but not a substitute for effective risk management (see also Gripaios, 1994). Following this line of thought, Fatemi and Glaum (2000) – for example – investigated the risk management practices of German non-financial firms regarding foreign exchange rate risk and interest rate risk and, unsurprisingly, concluded that forecasts are vital input factors in risk management.

Investigating banks from the US, UK, Germany, Canada and Japan, Madura and Zarruk (1995) delivered empirical evidence for significant interest rate exposure of these financial institutions. Furthermore, Spiwoks et al. (2008) emphasized the need for interest rate forecasts for banks to fulfill their tasks of maturity transformation.[20] Not only considering banks, Czaja et al. (2009) discussed the relevance of interest rate risk for German financial institutions. Schwarzbach et al. (2012), for example, annotated the relevance of ten year German government bond yield forecasts for life insurers' asset managers.[21]

19 See especially Parsley and Popper (2006) who convincingly emphasized the indirect effects of exchange rate movements on firms' profitability.

20 As noted above, financial institution also need forecasts to be compliant with regulatory requirements.

21 The authors motivated their investigations with the findings of Chopra and Ziemba (1993), who underlined the strong relevance of return forecasts for successful asset allocations.

In addition, governmental and monetary policy decision-makers have already been interested in the outlook for interest rates, oil price and exchange rate forecasts for decades (Elliott and Timmermann, 2008; Duffee, 2013; Butter and Jansen, 2013). Consequently, the relevance of expectations regarding financial market variables and financial market forecasts has also been intensively discussed from the perspective of these recipients.

For example, governmental decision-makers conduct fiscal policy not only in accordance with the current and expected future stance of the economy but also under simultaneous consideration of the current and expected state of financial markets. Regarding interest rates, the EMU sovereign debt crisis impressively demonstrated that market participants' expectations are strongly relevant concerning the refinancing needs of governments and the corresponding costs in terms of interest payments. Government bond yields have been substantially influenced by the changing assessment of investors (see, for example, De Santis, 2014, who emphasized the impact of flight-to-quality effects on government bond yields of crisis-burdened EMU members). Furthermore, one might especially argue that governmental decision-makers are particularly interested in the foreign exchange market as it affects their countries' economic competitiveness.[22] More recently, the impact of commodity price risk or commodity price shocks, respectively, on sovereigns' credit quality and governments' fiscal stability has been discussed intensively (see, for example, Van Der Ploeg, 2017; Lopez-Martin et al., 2017). Hence, especially for commodity exporting countries – that strongly rely on the receipts from processing and exporting their natural resources – commodity price forecasts and measures to indicate uncertainties regarding commodity prices hold strong relevance.

And finally, monetary policy decision-makers generally conduct actions in accordance with their mandate. For example, the European Central Bank (ECB) focuses on price stability, whereas the U.S. Federal Reserve sets its monetary policy in accordance with its goals of price stability, maximum employment, and moderate long-term interest rates (McCandless Jr and Weber, 1995; Pollard, 2003; Elliott and Timmermann, 2008; Basse et al., 2017). These mandates make the application of any kind of forecast or expectation formation inevitably necessary. Furthermore, due to the interdependent character regarding their reaction to and influence on financial markets central banks traditionally received a tremendous amount of attention when discussing financial market forecasts as well as market expectations (Blinder, 2000). In fact, central bankers are both recipients of forecasts for and generate impulses to financial markets. As one example, scholars have intensively discussed the possible reactions of central banks to exchange rate movements (Taylor, 2001; Lubik and Schorfheide, 2007).[23] Dealing with monetary policy decision-makers' reaction to financial market developments and uncertainties, Pagan and Robertson (2008) provided a comprehensive overview with regard to the relevant topics for central banks using forecasts as a decision support tool. Furthermore, market participants' expectations hold strong relevance for central banks (see, for example, Neuenkirch, 2012, who discussed the relationship between central bank communication and market expectations). Conversely, central bank communication may theoretically reduce uncertainty in the way that it increases forecast accuracy. In fact, it is argued here

[22] The relevance of exchange rates has been emphasized – for example – for the individual economies in the euro area (Breuer and Klose, 2015; Lucarelli et al., 2018). The political importance of FX markets also becomes clear in the current debate regarding the question of whether China is a "currency manipulator" (Ramirez, 2013).

[23] Reversely, and as discussed below, financial market participants' expectations are driven by central bank communication, its forward guidance and its actual measures.

that lower uncertainty has to be seen as one intended effect of central banks' forward guidance (see also Chortareas et al., 2002; Reeves and Sawicki, 2007; Trabelsi, 2016, who investigated the interrelationships of central bank transparency, forward guidance and forecasting from varying perspectives). This issue can also be understood as a nexus between central banks' reaction functions to (expected) future developments and the potential influence of their monetary policy alignment on financial markets. The latter especially finds expression via the interest rate channel, the exchange rate channel as well as the credit channel (Mishkin, 1995; Boivin et al., 2010). Moreover, regarding oil prices, Amendola et al. (2017) found empirical evidence indicating a positive relationship between expansionary monetary policy and oil price volatility. Due to this mutual influence, forecasters regularly have to build an understanding of the reaction function of central bankers (Sturm and De Haan, 2011) and monetary policy decision-makers have to anticipate economy wide second and third-round effects of their words and deeds.

In fact, financial market variables themselves regularly serve as predictors for economic activity and recessions. Unsurprisingly, a voluminous strand of literature has emerged around this topic. Following – for example – Harvey (1991), *"interest rates provide a window for future economic growth"* (see Harvey, 1991, p. 701). The shape of the yield curve has been frequently applied to predict economic downturns (Estrella and Mishkin, 1998; Ahrens, 2002; Wheelock and Wohar, 2009; Bluedorn et al., 2016). Following the discussion above, central bankers' interest regarding the outlook for crude oil prices may be initially associated with its influence on the macroeconomic price level. Having said that, the relationship between global crude oil prices and economic activity in general also holds special interest for monetary policy decision-makers (Barsky and Kilian, 2001; Kilian, 2009; Miao et al., 2017). For example, Bernanke et al. (1997) annotated that *"in the view of many economists, oil price shocks are perhaps the leading alternative to monetary policy as a key factor in postwar U.S. recessions"* (Bernanke et al., 1997, p. 93). Focusing on – but not limited to – the US economy Mohaddes and Pesaran (2017) recently investigated the effects of falling oil prices on interest rates, inflation and stock prices as well as dividends (used as a proxy for real economic activity). Regarding the effects on economic growth, the authors concluded that lower oil prices are in fact beneficial for real economic activity. However, the uncertainty of oil prices may have a negative effect on investment (Elder and Serletis, 2010) and thus might hinder economic activity. Following this line of thought, Jo (2014) discussed the impact of oil price uncertainty on global growth measured by industrial production and concluded that there exists an inverse relationship.[24]

Summarizing these thoughts it is reasonable to state that interest rates, crude oil prices as well as exchange rates – and especially forecasts thereof – are crucial variables for decision-makers. In this context, institutional backgrounds, changes in price-building processes and external impulses hold strong relevance. The three studies of this thesis presented in Chapters 2, 4 and 6 address recent and specific developments regarding government bond, crude oil and foreign exchange markets.

[24] Interestingly, the authors emphasized the need for survey-based measures for uncertainty regarding energy prices.

Survey forecasts for financial markets: evaluations, expectations and uncertainty

A wide range of forecasting methods exist that are applied to predict the future movements of economic indicators and financial variables.[25] As noted above, this thesis addresses the applications and evaluations of survey forecasts for financial market variables.[26] Especially for decision-makers who are not able or willing to produce their own forecasts for all relevant or required financial or economic variables survey forecasts may be important input variables. In this context, survey forecasts might also be seen as a potential alternative for self-produced predictions – or may at least be used as a means of benchmarking.

The evaluation of forecasts and prediction models is crucial for policy decision-makers (Dovern and Ziegler, 2008). For recipients of survey forecasts it is important to assess and compare the predictions received using reasonable and easily reproducible quality criteria. Forecast accuracy measures are regularly applied as a standard of comparison for two or more competing forecasting approaches. Commonly-deployed metrics regarding accuracy are statistical error measures[27] as well as measures of sign accuracy.[28] In addition, measures of relative forecast accuracy have been regularly applied. Here, the Theil's U (Theil, 1955; Theil, 1992) as well as the Diebold Mariano test of equal predictive accuracy (Diebold and Mariano, 1995) are widely used. Testing empirically for the presence of a status quo bias, Andres and Spiwoks (1999) introduced the TOTA (topically oriented trend adjustment) coefficient.[29]

Over and above these accuracy measures, scholars have intensively investigated the additional information that can be delivered by survey forecasts and a vast body of literature is dedicated to the concept of forecast combination.[30] For example, Kim and Orphanides (2012) successfully utilized the additional information of survey forecasts when modeling the term structure of interest rates. However, given the focus of this thesis, two interconnected applications of survey forecasts are outstanding and will be discussed in further detail below: first, survey forecasts may be applied to investigate the expectation formation of market participants and hence may aid in inspecting price-building-processes; and second, the heterogeneity of survey

[25] There does not exist a distinct and universally-valid classification. Good overviews and discussions of different applied forecasting methodologies can be found in Chatfield (1997), Fauvel et al. (1999), Hendry and Clements (2003), Fildes et al. (2008) as well as De Gooijer and Hyndman (2006). Moreover, focusing on major central banks, Butter and Jansen (2013) highlighted that the applied forecasting methods are in fact quite diverse, ranging from pure judgment and expert advisory to (complex) macro-econometric models.

[26] It is important to note that the investigated surveys – *Consensus Economics* (see Chapter 3 and Chapter 7) as well as *Dow Jones* (see Chapter 5) – are collected from professional forecasters. Although the explicit forecasting methodologies applied by these experts are not known, it is reasonable to assume that the forecasters' approaches show considerable differences and may range from judgmental forecasts to more complex econometric models (see Pesaran and Weale, 2006, who also provide a comprehensive overview of further collections of survey predictions).

[27] The root mean squared error belongs to the most frequently-applied statistical error measures (see, for example, Kolb and Stekler, 1996).

[28] See, for example, Greer (2003) who applied the sign accuracy test when evaluating interest rate forecasts.

[29] The TOTA coefficient has been frequently applied for interest rates (see, for example, Spiwoks and Hein, 2007; Spiwoks et al., 2010) and exchange rates (see, for example, Bofinger and Schmidt, 2003).

[30] See Bates and Granger (1969), Granger and Ramanathan (1984), Clemen (1989), Hendry and Clements (2004), Greer (2005), Clements and Harvey (2011), and Blanc and Setzer (2016).

predictions might deliver relevant and valuable insights into market participants' uncertainty.

Following Pesaran and Weale (2006), the formation of expectations is an *"integral part of the decision-making process by households, firms, as well as the private and public institutions"* (see Pesaran and Weale, 2006, p. 717.) and *"expectations are subjectively held beliefs by individuals about uncertain future outcomes or the beliefs of other individuals in the market place"* (see Pesaran and Weale, 2006, p. 720.). Following – for example – Ruelke et al. (2012), the assumption of traders' rational forecasts is a *"cornerstone"* in capital market theory (see Ruelke et al., 2012, p. 2757). When used as a proxy for market expectations, one major advantage of survey forecasts stems from the fact that they are exogenous (MacDonald and Torrance, 1988; Jongen et al., 2008).[31] Consequently, the rational expectation hypothesis brought forth by Muth (1961) received strong attention in the context of investigating financial market survey data (MacDonald, 2000).[32]

For the three financial markets in the focus of this thesis, survey forecasts have been utilized for decades.[33] After surveying a substantial part of the evaluation literature, MacDonald (2000) concluded that biasedness and inefficiency of survey forecasts are common findings.[34] However, he also mentioned the capability of survey forecasts to investigate the behavior of asset markets also taking risk premia in financial markets into consideration.

Moreover, in the context of the empirical investigations of survey forecasts in Chapters 3 and 7 of this thesis, some methodological remarks are reasonable. In order to assess the rationality of the financial forecasts (i.e. interest rates in Chapter 3 and exchange rates in Chapter 7) a focus will be placed on alternative measures that are located in the field of applied time series analysis, following the reasoning of Cheung and Chinn (1998). The authors proposed specifically investigating the statistical properties of the relevant time series. In this sense, as a necessary condition for rationality, the time series of the actual observations and the corresponding forecasts should share the same order of integration and should be cointegrated (Liu and Maddala, 1992; Cheung and Chinn, 1998).[35] Provided that both preconditions are met further investigations with regard to the forerunning properties of forecasts are feasible. It is argued in this thesis that *good forecasts* should fulfill the requirement of rationality following Cheung and Chinn, 1998 and deliver a forward-looking view.

[31] Beckmann and Czudaj (2017a) recently annotated that consensus forecasts are *"considered to be the most adequate approximation of market expectations available"* (see Beckmann and Czudaj, 2017a, p. 149).

[32] Common tests regarding the rationality of survey forecasts for financial market variables are the test for unbiasedness and the test for forecast efficiency; also known as test for orthogonality of error terms (Nordhaus, 1987; Ito, 1990; Ronald and Ian, 1993).

[33] Pioneering studies come – for example – from Friedman (1980) focusing on interest rates, Dominguez (1986) investigating survey predictions for exchange rates and Ronald and Ian (1993) examining crude oil forecasts.

[34] More than a decade ago, MacDonald (2000) investigated the evaluation literature for bond markets, foreign exchange markets and stock markets. More recent studies also dealing with oil prices confirm the findings of MacDonald (2000). For oil price studies, see, for example, Ronald and Ian (1993), Reitz et al. (2009), as well as Prat and Uctum (2011). Chortareas et al. (2012) focused on interest rate forecasts. Surveying the literature on exchange rate surveys Jongen et al. (2008) came to similar conclusions like MacDonald (2000). More recent results regarding exchange rates can be found in Beckmann and Czudaj (2017a) and Ince and Molodtsova (2017). For example, Ince and Molodtsova (2017) discussed how survey forecasts for exchange rates allow testing the assumption of rational expectations of market participants and inter alia could not accept the hypothesis of unbiasedness.

[35] Two time series are said to be cointegrated when they share a common stochastic trend (See, for example, Granger, 1981; Engle and Granger, 1987; Lütkepohl and Krätzig, 2004).

The latter may be statistically validated using Granger causality analysis.[36] Added together this framework offers a straightforward and structural assessment of financial market forecasts. However, in the empirical literature comparatively few studies utilize this approach and hence the studies in Chapters 3 and 7 provide novel and relevant contributions to the field of forecast evaluation.[37]

When analyzing survey data, it becomes obvious that from a cross sectional forecaster perspective individuals' expectations may vary substantially (see also Mankiw et al., 2003; Sill, 2014). Furthermore, it is important to bear in mind that when investigating expectation formation coping with uncertainty is unavoidable and relevant (see also Poncela and Senra, 2017). In this context, it is crucial to clearly differentiate between risk and uncertainty, as Knight (1921) expressed in his seminal contribution. Following Knight (1921), risk is quantifiable with known probabilities and for uncertainty there are no known probabilities of events (see Knight, 1921; Basili, 2001). The relevance of this *Knightian uncertainty* for financial markets has been emphasized in the literature (Dow and Costa Werlang, 1992; Epstein and Wang, 1994; Basili, 2001; Rigotti and Shannon, 2005). For example, Basili (2001) annotated that an *"uncertainty attitude of agents may shed new light on some financial market puzzles and provides an new explanation of them."* (Basili, 2001, p. 2).

Following this line of thought, survey forecasts may provide a missing link not only to test the rationality of market participants but also to measure uncertainty (Lahiri and Sheng, 2010). Poncela and Senra (2017), listed three approaches applied to measure uncertainty using survey forecasts. The authors mentioned disagreement from point forecasts, the equal weighted average of individual uncertainty and the equal weighted aggregation of the individual density forecasts (Poncela and Senra, 2017).[38] The empirical evidence regarding the usefulness of survey data in delivering uncertainty measures is rather mixed, whereas the majority of studies focus on macroeconomic indicators (see also Atalla et al., 2016). Starting with Zarnowitz and Lambros (1987), scholars have been discussing – and questioning – the usefulness of forecasters' disagreement as a gage for uncertainty (for comprehensive overviews of the relevant literature see Boero et al., 2008; Lahiri and Sheng, 2010; Abel et al., 2016). In fact, Jurado et al. (2015) listed relevant drawbacks when using survey forecasts in this context.[39]

However, it is argued in this thesis that given the limited availability of alternative measures for uncertainty as well as the relevance of financial market variables for economic activity, survey forecasts should not be ruled out in general. In fact, survey forecasts for financial market variables may be especially useful to draw inference in the context of the global financial crisis. This has been documented by Beckmann and Czudaj (2017b), who were able to demonstrate a close link between

[36] Generally speaking, a time series Granger causes another time series when past values of the former deliver additional informational content for the prediction of the latter (Engle and Granger, 1987; Gelper and Croux, 2007).

[37] Some examples in the forecast evaluation literature are Berk (1999) (focusing on inflation forecasts), Schwarzbach et al. (2012) as well as Kunze et al. (2014) (focusing on interest rate forecasts) and the recent study of Cheung et al. (2017a) (focusing on exchange rates).

[38] In fact, the availability of surveys providing the analysts' individual probability distribution are rather scarce. Hence, the majority of studies evaluates disagreement among point forecasts (see also Abel et al., 2016).

[39] The authors highlighted among others the limited number of variables with corresponding forecasts as well as issues with behavioral biases of forecasters.

forecasters' disagreement and uncertainty. Additionally, disagreement among forecasters used as dispersion measures may aid explaining the financial market participants' expectation formation process (Dovern, 2015).

In accordance with these lines of thought and focusing on interest rates Kunze et al. (2014) used the high low spread of monthly forecasts for the three months interbank rate provided by *Consensus Economics* as a measure for financial market uncertainty and successfully explained crisis-related movements of this variable with financial stress indicators and economic sentiment measures. Focusing on survey forecasts for exchange rates, Cavusoglu and Neveu (2015) recently emphasized the predictive power of dispersion measures. Regarding oil prices, utilizing data collected by *Consensus Economics* Singleton (2013) was among the first to draw on dispersion measures derived from survey forecasts, albeit not focusing on uncertainty per se. Also focusing on oil prices, Atalla et al. (2016) investigated the dispersion of forecasts contributing to the ECB's *Survey of Professional Forecasters* and found empirical evidence that rising forecasters' disagreement mirrors increased uncertainty. They also concluded that oil price volatility can explain forecasters' disagreement. The three studies in this thesis focusing on the evaluation of survey forecasts take on the concepts discussed in this section, albeit in varying degrees (see Chapters 3, 5 and 7.).

Summary of the studies on interest rates and interest rate forecasts

The first two studies of this thesis presented in Chapters 2 and 3 deal with crisis-related effects on EMU interest rates as well as interest rate forecasts in Germany and the United Kingdom. The global financial crisis as well as the EMU sovereign debt crisis strongly influenced both short-term rates and the long end of the yield curve (see, for example, Basse, 2014; Kunze, 2014; Kunze et al., 2015). Especially in the case of long-term rates the comeback of risk premia demanded by government bond investors to hold sovereign debt as well as flight-to-quality effects became evident in the course of the financial crisis (see, for example, Kunze, 2014; Kunze et al., 2015). Despite being technically ruled out in a currency union, redenomination risks have been recently discussed for EMU government bonds (Klose and Weigert, 2014; Sibbertsen et al., 2014). In this context, rising yield spreads of different EMU government bonds have to be interpreted against the background of dramatically-altered market expectations regarding the fiscal sustainability of EMU member countries and rising uncertainty. This is also true regarding the outlook for short-term interest rates. In fact, significantly-altered monetary policy alignments of the central banks in Washington, Frankfurt and London applied to counter economic downturns have to be seen as a crisis related impulses on short term rates and interbank rates (see, for example, Kunze et al., 2014).

Since survey forecasts are frequently used as exogenously given proxies for expectations, forecast errors as well as measures of forecast dispersion have recently received a substantial amount of attention. Furthermore, as discussed above, it has been convincingly argued in the literature that interest rates – as financial indicators – may be applied to predict real economic activity or recessions, respectively (Fisher, 1907; Harvey, 1991; Estrella and Mishkin, 1998). The studies in Chapters 2 and 3 address these issues and deliver an empirical investigation of crisis-related changes to EMU government bond yields from an asset manager's perspective, investigate the accuracy and rationality of interest forecasts and link dispersion measures of survey forecasts to real economic indicators focusing on the implications for (policy) decision-makers.

In Chapter 2 – *Asset liability management and the euro crisis: Sovereign credit risk as a challenge for the German life insurance industry* – the incisive adjustment regarding market participants' attitude towards non-core EMU government debt is investigated. Hence, it is firstly argued that the rise in government bond yield spreads between Germany and crisis-burdened EMU member countries has to be associated with a break in a long run relationship of the country-specific long term government bond yields. The second objective of our investigation is the real life application of this far-reaching development for asset and risk managers in the financial services industry. Thereby we focus on the German life insurance industry, given that at least in comparison with the large body of research dealing with effects of the EMU sovereign debt crisis on the banking industry, the literature focusing on insurance companies in this regard is much less extensive (Düll et al., 2017). Taking on this literature gap, Chapter 2 begins by providing a detailed overview of the investment patterns of German life insurers and discusses the relevant impacts of the EMU sovereign debt crisis.

We particularly illustrate that, given the large share of interest-baring assets, the interest rate level holds strong significance for the life insurance industry. Especially asset managers in the German life insurance industry seem to have had a complex task. German government bond yields touched record lows in the aftermath of the crisis, due to safe haven effects, deflationary fears as well as the ECB's monetary policy adjustments. Hence, relying solely on German government bonds would have had severe implications for life insurers' interest income. However, higher yielding assets issued by Spain and Italy seemed to be much riskier, due to sovereign credit risk and redenomination risk. Unsurprisingly, the EMU sovereign debt crisis strongly influenced the asset values of the German life insurance industry. We investigate the yield spreads of ten year government bond yields between Germany and Italy and Germany and Spain (viewing Italy and Spain as EMU non-core countries) as well as Germany and the Netherlands and Germany and Austria (viewing the Netherlands and Austria as EMU core countries) using unit root breakpoint tests (see Perron, 1990; Perron and Vogelsang, 1992; Vogelsang and Perron, 1998; Perron, 2006). One clear advantage of this straightforward test procedure stems from the fact that it allows checking for breakpoints with unknown timings. The statistical evidence from the applied unit root breakpoint tests confirms previous findings regarding euro crisis-related structural shifts in the cointegrating relationship between German and Spanish and German and Italian government bond yields, respectively. Hence, based on our statistical results we have been able to show that sovereign credit and probably redenomination risks have led to structural changes in the relationship between government bond yields in peripheral countries and Germany.

It is well documented in the empirical literature that from the perspective of decision-makers, measures of forecast accuracy and tests for rationality are reasonable complements in forecast evaluation (Dovern and Weisser, 2011; Chen et al., 2016; Ince and Molodtsova, 2017). Following this reasoning, Chapter 3 of this thesis – *Forecasting European interest rates in times of financial crisis – What insights do we get from international survey forecasts?* – presents an evaluation of survey forecasts provided by *Consensus Economics* utilizing both accuracy measures and tests for the rationality of forecasts. Given the structure of our data set we are able to focus on alternative tests for the rationality of the survey forecasts (our framework is based on the approach proposed in Cheung and Chinn, 1998).

Moreover, and in order to derive additional potential utilities from the survey forecasts, we examine possible implications from dispersion measures and forecast errors. Testing for structural changes we are able to detect mean shifts in the corresponding time series of these uncertainty measures. In fact, the timing of these breakpoints underpins the viewpoint that dispersion measures and forecast errors are suitable indicators for uncertainty. Empirical evidence in this regard is rather scarce in the academic literature. Moreover, based on Granger causality analysis, we provide new empirical evidence for the existence of a link between uncertainty regarding future interest rates and real economic activity. This finding should prove especially useful for fiscal or monetary policy decision-makers and financial market participants. However, more work has to be conducted in this growing research field. Especially in the context of early warning indicators or early warning systems, this approach might potentially deliver valuable insights, but has to be extended in terms of investigated countries as well as economic and financial indicators.[40]

Summary of the studies on crude oil prices and crude oil forecasts

The following two studies of this thesis presented in Chapters 4 and 5 are particularly concerned with the impact of oil prices on market expectations regarding sovereign credit risk of oil-exporting countries and the appraisal of oil price forecasting in the context of managerial decision-making. It has been intensively discussed in the academic literature that decision-makers of various kinds have to cope with the uncertainty of global crude oil prices. Empirical evidence indicates that the financing of some oil-exporting countries' budgets strongly relies on oil revenues. Hence, policy decision-makers in oil exporting countries have to plan their budget considering assumptions of future oil prices. As the EMU sovereign debt crisis impressively demonstrated, the sustainability of public finances has a substantial impact of market expectation regarding credit risk. In this context, the findings of Cimadomo et al. (2016) are highly relevant. The authors investigated survey forecasts for ten year government bond yields inter alia in Germany, Italy and France and highlighted that better fiscal outlooks induce lower forecasts regarding sovereign spreads.

Contributing to this strand of research, the third study in this thesis presented in Chapter 4 named *Oil prices and sovereign credit risk of oil producing countries: an empirical investigation* delivers novel empirical evidence on the relationship between sovereign credit risk and crude oil prices. We investigate daily observations of credit default swap spreads of eight oil-producing countries[41] and the price of crude

[40] Recently, Istrefi and Mouabbi (2017) utilized dispersions of *Consensus Economics* survey forecasts for short- and long-term interest rates in the US, Japan, Germany, France, the UK, Italy, Canada, Spain and Sweden. The authors argued that interest rate uncertainty may be attributed to uncertainty regarding monetary policy and hence has negative impacts on economic activity (Istrefi and Mouabbi, 2017).

[41] The countries in our data sample are the UK, US, Norway, Saudi Arabia, Brazil, Malaysia, Russia, Qatar and Venezuela.

oil using methods of advanced time series analysis. Given the non-stationarity of the time series and the empirical evidence pointing to generalized autoregressive conditional heteroskedasticity (GARCH), we employ a vector autoregressive VAR(1)-GARCH(1,1)-in-mean model. With this model, we are able to draw inference about the impact of crude oil price changes on the market perception of sovereign credit risk. Using this methodological approach to investigate the relationship between sovereign CDS spreads and crude oil prices delivers several noteworthy and relevant results. Most importantly, we have been able to document an inverse relationship between crude oil prices and sovereign CDS spreads for Saudi Arabia, Brazil, Malaysia, Russia, Qatar and Venezuela. These results indicate that market participants interpret rising oil prices as a positive influence on the fiscal stability of oil-producing countries. Interestingly, for Norway, the UK and the US, we did not find a significant inverse relationship and attribute these results to the diversified character of these economies. Given the relevance of sovereign credit risk, our recognitions hold strong impact for financial market practitioners and political decision-makers[42] and could also be seen as a starting point for further empirical research. For example, Shahzad et al. (2017) highlighted the relevance of oil volatility shocks for directional forecasts for sovereign credit risk (for more recent studies dealing with the relationship of oil price shocks and sovereign credit risk, see also Bouri et al., 2018; Lee and Lee, 2018).

Under the assumption that from a decision-making perspective it is not an option to abandon forecasts per se, oil price surveys may be a welcome input factor for decision-makers at governments, central banks, companies as well as financial institutions. In Chapter 5 entitled *The usefulness of oil price forecasts – evidence from survey predictions* we deliver evaluation results and appraise them in the context of managerial decision-making. To the best of our knowledge, the second oil price study presented in Chapter 5 utilizes for the first time the *Dow Jones Oil Price Survey*, which covers the two globally most important crude oil benchmarks. Hence, this contribution delivers novel empirical evidence regarding the evaluation of oil price forecasts. One major advantage of the data sample used stems from the fact that each month twelve forecast horizons are available. This allows drawing inference regarding the relationship between forecast quality and the forecast horizon. We find empirical evidence for status quo-dependent forecasts. Using the widely-acknowledged TOTA-coefficient, we are able to show that topically-oriented trend adjustment is even more pronounced for longer forecast horizons. However, for longer horizons forecasters outperform the naïve prediction. Regarding the rationality of forecasts, longer term forecasts seem to be unbiased. Having said that, the range of forecast errors – measured by the mean error and the adjusted root mean squared error – rises with the forecast horizon. This can be seen as evidence of higher uncertainty. Comparing these evaluation results with empirical findings regarding the quality of interest rate or exchange rate forecasts it has to be concluded, that oil price forecasts do not stand out against predictions for “purely financial” markets.

[42] Nusair (2016) investigated economic activity in Gulf Cooperation Council (GCC) countries and concluded that rising oil prices have a stronger impact on real GDP growth than falling oil prices. Interestingly, Nusair (2016) emphasized the need to diversify the GCC economies, which to a certain extent corresponds to our findings for the US, UK and Norway regarding financial markets’ perception of sovereign credit risk.

Summary of the studies on exchange rates and exhange rate forecasts

The last two studies of this thesis presented in Chapters 6 and 7 focus on the foreign exchange markets in Asia. In general, exchange rates have a significant impact on foreign trade and cross border investments (see, for example, Kan, 2017) and are important factors for open economies and international finance (see Dreger and Stadtmann, 2008; Dick et al., 2015). First of all, we motivate our research focus on Asia based on this region's rising share of global GDP growth, cross border investments and world trade. Second, given the heterogeneous landscape regarding its exchange rate markets and regimes, the Asian hemisphere holds special interest for researchers as well as political and managerial decision-makers.[43] Moreover, the most significant economic driving force in that region is the Chinese economy, which – despite its ongoing protectionist measures – becomes increasingly integrated into global value chains and – at least gradually – into financial markets. Hence, Asia in general and China in particular obviously offer themselves as research objectives regarding institutional changes in foreign exchange markets.

In Chapter 6 – *The global emergence of the RMB: A "New Normal" for China's FX markets?* – we investigate the current state of Renminbi (RMB) internationalization and the characteristics of China's bipolar exchange rate system. For decades, researchers investigated China's growth miracle, its international interdependencies and its protectionist measures (Ma and McCauley, 2008; Prasad, 2009; Morrison, 2011; Funke et al., 2015). Especially, the exchange rate policy has been criticized (Frankel and Wei, 2007; Hu et al., 2016). Pegging the RMB to the US dollar has been considered a substantial growth driver for the Chinese economy, which comes at the cost of China's trade partners (see Goldstein and Lardy, 2006). Meanwhile, this debate's focus shifted to RMB internationalization (Batten and Szilagyi, 2016; Cui, 2017).[44] Furthermore, it has been discussed in the literature whether the RMB already plays a dominant role in Asia (Subramanian and Kessler, 2013), whether there will be a transition to a tripolar currency system (see, for example, Fratzscher and Mehl, 2014) or whether the RMB may eventually succeed the US dollar as the global anchor currency (see, for example, Ito, 2010).

Having said that, the RMB is not yet freely tradable and still subject to a large degree of intervention by China's monetary policy decision-makers. Focusing on the bipolar structure of China's FX market – which builds on the parallel existence of a mainland market and offshore trading locations – we argue in Chapter 6 that despite ongoing internationalization efforts the imperfect integration of on- and offshore markets for the Chinese currency is a substantial impediment for RMB internationalization.[45]

43 From a political perspective especially the vast foreign exchange reserves in Asia as well as ongoing foreign exchange interventions have been intensively discussed in the literature (Aizenman and Lee, 2008; Pontines and Rajan, 2011; Ouyang and Rajan, 2011).

44 In this context the inclusion of the RMB into the IMF's basket of special drawing rights (Dixon et al., 2016) and the growing relevance in terms of becoming an international payment currency (Cheung and Rime, 2014; Zhang and Zhang, 2017) are important developments.

45 In fact, it has been discussed in the literature that the so-called CNH-CNY-spread has been affected by arbitrage and carry trades (Liu, 2015; Zhang and Zhang, 2017) as well as global risk aversion. Funke et al. (2015) and Cheung et al. (2017b) cited so called "risk on, risk off" cycles.

We investigate market integration between on- and offshore markets using methods of advanced time series analysis and find empirical evidence for strong persistence of the spread between the on- and offshore quotations for the RMB. Furthermore, we do not find any empirical evidence that the degree of market integration has improved lately. We infer from these results that the Chinese FX market is far from being perfectly integrated and hence has not reached its "New Normal" yet.

In the last chapter of this thesis headlined *Predicting exchange rates in Asia: New insights on the accuracy of survey forecasts* aggregated survey forecasts provided by *Consensus Economics* for the exchange rates of the Chinese yuan, the Hong Kong dollar, the Japanese yen, and the Singapore dollar vis-à-vis the US dollar are evaluated. Using common forecast accuracy measures it is shown that all forecasts investigated are irrational in the sense that the predictions are biased. However, these results are inconsistent with the aforementioned alternative measure of rationality based on methods of applied time series analysis. Investigating the order of integration of the time series and using cointegration analysis, empirical evidence supports the conclusion that the majority of forecasts are rational. However, the forerunning properties of the predictions are less convincing. One major contribution of the study to the literature of forecast evaluation is its focus on different currency regimes. In this regard, a comprehensive up-to-date overview of the International Monetary *de facto* exchange rate regimes for China, Hong Kong, Japan and Singapore is provided. In fact, the investigated foreign exchange rate regimes range from free floating (JPY) to currency board systems (HKD). Empirical evidence indicates that the currency regime matters for the quality of exchange rate forecasts. Regarding impulses for further research, the study could also be seen as starting point for additional investigations of regime-dependent investigations of FX forecasts. Especially regarding event risks (for example, shifts in the FX regimes), further research in the field of forecast evaluation – also incorporating disaggregated survey data – is necessary.

Focusing on interest rates, oil prices and exchange rates this thesis delivers new insights on price-building processes, market expectations and predictability for these three financial markets. Furthermore it presents novel evidence regarding the accuracy, rationality and usefulness of survey based interest rate, crude oil price and exchange rate forecasts. Given the indispensability of forward-looking planning these appraisals of available forecasts still are crucial for decision-makers. Furthermore, it is shown that survey forecasts offer useful additional insights regarding forecasters' rationality and uncertainty.

References

Abel, Joshua, Robert Rich, Joseph Song, and Joseph Tracy (2016). "The measurement and behavior of uncertainty: Evidence from the ECB survey of professional forecasters". In: *Journal of Applied Econometrics* 31.3, pp. 533–550.

Ahrens, Ralf (2002). "Predicting recessions with interest rate spreads: A multicountry regime-switching analysis". In: *Journal of International Money and Finance* 21.4, pp. 519–537.

Aizenman, Joshua and Jaewoo Lee (2008). "Financial versus monetary mercantilism: Long-run view of large international reserves hoarding". In: *The World Economy* 31.5, pp. 593–611.

Alquist, Ron, Lutz Kilian, and Robert J Vigfusson (2013). "Forecasting the price of oil". In: *Handbook of Economic Forecasting*. Ed. by G. Elliott, C.W.J. Granger, and A. Timmermann. Vol. 2. Elsevier, pp. 427–507.

Amendola, Alessandra, Vincenzo Candila, and Antonio Scognamillo (2017). "On the influence of U.S. monetary policy on crude oil price volatility". In: *Empirical Economics* 52.1, pp. 155–178.

Andres, Peter and Markus Spiwoks (1999). "Forecast quality matrix: A methodological survey of judging forecast quality of capital market forecasts". In: *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)* 219.5+6, pp. 513–542.

Aretz, Kevin, Söhnke M Bartram, and Gunter Dufey (2007). "Why hedge? Rationales for corporate hedging and value implications". In: *The Journal of Risk Finance* 8.5, pp. 434–449.

Atalla, Tarek, Fred Joutz, and Axel Pierru (2016). "Does disagreement among oil price forecasters reflect volatility? Evidence from the ECB surveys". In: *International Journal of Forecasting* 32.4, pp. 1178–1192.

Baghestani, Hamid, Mohammad Arzaghi, and Ilker Kaya (2015). "On the accuracy of Blue Chip forecasts of interest rates and country risk premiums". In: *Applied Economics* 47.2, pp. 113–122.

Barron, Michael and David Targett (1988). "Managing forecasts: User and expert". In: *Management Decision* 26.1, pp. 41–44.

Barsky, Robert B and Lutz Kilian (2001). "Do we really know that oil caused the great stagflation? A monetary alternative". In: *NBER Macroeconomics annual* 16, pp. 137–183.

Bartram, Söhnke (2002). "The interest rate exposure of non-financial corporations". In: *Review of Finance* 6.1, pp. 101–125.

Bartram, Söhnke M, Gregory W Brown, and Bernadette A Minton (2010). "Resolving the exposure puzzle: The many facets of exchange rate exposure". In: *Journal of Financial Economics* 95.2, pp. 148–173.

Basili, Marcello (2001). "Knightian uncertainty in financial markets: An assessment". In: *Economic Notes* 30.1, pp. 1–26.

Basse, Tobias (2014). "Searching for the EMU core member countries". In: *European Journal of Political Economy* 34, S32–S39.

Basse, Tobias, Frederik Kunze, Bernd Krampen, and Christoph Schwarzbach (2017). "German bond markets and U.S. monetary policy – Some thoughts from the perspective of the European life insurance industry". In: *Zeitschrift für die gesamte Versicherungswissenschaft*, pp. 1–12.

Bates, JM and CWJ Granger (1969). "The combination of forecasts". In: *Journal of the Operational Research Society* 20.4, pp. 451–468.

Batten, Jonathan A. and Peter G. Szilagyi (2016). "The internationalisation of the RMB: New starts, jumps and tipping points". In: *Emerging Markets Review* 28, pp. 221–238. ISSN: 1566-0141.

Beckmann, Joscha and Robert Czudaj (2017a). "Exchange rate expectations since the financial crisis: Performance evaluation and the role of monetary policy and safe haven". In: *Journal of International Money and Finance* 74, pp. 283–300.

— (2017b). "The impact of uncertainty on professional exchange rate forecasts". In: *Journal of International Money and Finance* 73, pp. 296–316.

Bekaert, Geert, Marie Hoerova, and Marco Lo Duca (2013). "Risk, uncertainty and monetary policy". In: *Journal of Monetary Economics* 60.7, pp. 771–788.

Belongia, Michael (1987). "Predicting interest rates: A comparison of professional and market-based forecasts". In: *Federal Reserve Bank of St. Louis Review* Mar, pp. 9–15.

Berk, Jan Marc (1999). "Measuring inflation expectations: A survey data approach". In: *Applied Economics* 31.11, pp. 1467–1480.

Bernanke, Ben S, Mark Gertler, Mark Watson, Christopher A Sims, and Benjamin M Friedman (1997). "Systematic monetary policy and the effects of oil price shocks". In: *Brookings Papers on Economic Activity* 1997.1, pp. 91–157.

Blanc, Sebastian M and Thomas Setzer (2016). "When to choose the simple average in forecast combination". In: *Journal of Business Research* 69.10, pp. 3951–3962.

Blinder, Alan S (2000). "Central-bank credibility: Why do we care? How do we build it?" In: *American Economic Review* 90.5, pp. 1421–1431.

Bloom, Nicholas (2009). "The impact of uncertainty shocks". In: *Econometrica* 77.3, pp. 623–685.

Bluedorn, John C, Jörg Decressin, and Marco E Terrones (2016). "Do asset price drops foreshadow recessions?" In: *International Journal of Forecasting* 32.2, pp. 518–526.

Bodnar, Gordon M and William M Gentry (1993). "Exchange rate exposure and industry characteristics: Evidence from Canada, Japan, and the USA". In: *Journal of International Money and Finance* 12.1, pp. 29–45.

Boero, Gianna, Jeremy Smith, and Kenneth F Wallis (2008). "Uncertainty and disagreement in economic prediction: the Bank of England survey of External Forecasters". In: *The Economic Journal* 118.530, pp. 1107–1127.

Bofinger, Peter and Robert Schmidt (2003). "On the reliability of professional exchange rate forecasts". In: *Financial Markets and Portfolio Management* 17.4, pp. 437–449.

Boivin, Jean, Michael T Kiley, and Frederic S Mishkin (2010). *How has the monetary transmission mechanism evolved over time?* Tech. rep. National Bureau of Economic Research.

Borio, Claudio (2014). "The financial cycle and macroeconomics: What have we learnt?" In: *Journal of Banking & Finance* 45, pp. 182–198.

Bouri, Elie, Syed Jawad Hussain Shahzad, Naveed Raza, and David Roubaud (2018). "Oil volatility and sovereign risk of BRICS". In: *Energy Economics*.

Breuer, Sebastian and Jens Klose (2015). "Who gains from nominal devaluation? An empirical assessment of Euro-area exports and imports". In: *The World Economy* 38.12, pp. 1966–1989.

Butter, Frank AG den and Pieter W Jansen (2013). "Beating the random walk: A performance assessment of long-term interest rate forecasts". In: *Applied Financial Economics* 23.9, pp. 749–765.

Caldara, Dario, Cristina Fuentes-Albero, Simon Gilchrist, and Egon Zakrajšek (2016). "The macroeconomic impact of financial and uncertainty shocks". In: *European Economic Review* 88, pp. 185–207.

Cardarelli, Roberto, Selim Elekdag, and Subir Lall (2011). "Financial stress and economic contractions". In: *Journal of Financial Stability* 7.2, pp. 78–97.

Cavusoglu, Nevin and Andre R Neveu (2015). "The predictive power of survey-based exchange rate forecasts: Is there a role for dispersion?" In: *Journal of Forecasting* 34.5, pp. 337–353.

Chatfield, Chris (1997). "Forecasting in the 1990s". In: *Journal of the Royal Statistical Society: Series D (The Statistician)* 46.4, pp. 461–473.

Chen, Qiwei, Mauro Costantini, and Bruno Deschamps (2016). "How accurate are professional forecasts in Asia? Evidence from ten countries". In: *International Journal of Forecasting* 32.1, pp. 154–167.

Cheung, Yin-Wong, Menzie D Chinn, Antonio Garcia Pascual, and Yi Zhang (2017a). *Exchange rate prediction redux: New models, new data, new currencies*. Tech. rep. National Bureau of Economic Research.

Cheung, Yin-Wong, Cho-Hoi Hui, and Andrew Tsang (2017b). "The Renminbi central parity: An empirical investigation". In: *BOFIT Discussion Papers No. 07/2017*.

Cheung, Yin-Wong and Dagfinn Rime (2014). "The offshore Renminbi exchange rate: Microstructure and links to the onshore market". In: *Journal of International Money and Finance* 49, pp. 170–189.

Cheung, Y-W and Menzie David Chinn (1998). "Integration, cointegration and the forecast consistency of structural exchange rate models". In: *Journal of International Money and Finance* 17.5, pp. 813–830.

Chopra, Vijay Kumar and William T Ziemba (1993). "The effect of errors in means, variances, and covariances on optimal portfolio choice". In: *The Journal of Portfolio Management* 19.2, pp. 6–11.

Chortareas, Georgios, Boonlert Jitmaneeroj, and Andrew Wood (2012). "Forecast rationality and monetary policy frameworks: Evidence from UK interest rate forecasts". In: *Journal of International Financial Markets, Institutions and Money* 22.1, pp. 209–231.

Chortareas, Georgios, David Stasavage, and Gabriel Sterne (2002). "Does it pay to be transparent? International evidence from central bank forecasts". In: *Review-Federal Reserve Bank of Saint Louis* 84.4, pp. 99–118.

Cimadomo, Jacopo, Peter Claeys, and Marcos Poplawski-Ribeiro (2016). "How do experts forecast sovereign spreads?" In: *European Economic Review* 87, pp. 216–235.

Clemen, Robert T (1989). "Combining forecasts: A review and annotated bibliography". In: *International Journal of Forecasting* 5.4, pp. 559–583.

Clements, Michael P and David I Harvey (2011). "Combining probability forecasts". In: *International Journal of Forecasting* 27.2, pp. 208–223.

Clements, Michael P and David F Hendry (2006). "Forecasting with breaks". In: *Handbook of Economic Forecasting* 1. Ed. by G. Elliott, C.W.J. Granger, and A. Timmermann, pp. 605–657.

Cox, Garfield V (1930). "Evaluation of economic forecasts". In: *Journal of the American Statistical Association* 25.169A, pp. 31–35.

Creal, Drew D and Jing Cynthia Wu (2017). "Monetary policy uncertainty and economic fluctuations". In: *International Economic Review* 58.4, pp. 1317–1354.

Cui, Yuming (2017). "The internationalization of the RMB: A perspective vis a vis East Asian economic and financial integration". In: *Asia Pacific Business Review*, pp. 1–19.

Czaja, Marc-Gregor, Hendrik Scholz, and Marco Wilkens (2009). "Interest rate risk of German financial institutions: the impact of level, slope, and curvature of the term structure". In: *Review of Quantitative Finance and Accounting* 33.1, pp. 1–26.

Damie, C, C Jeffrey, B Abhisek, O Raquel, and H Michael (2015). "The new oil order". In: *The Goldman Sachs Group, Inc*.

De Gooijer, Jan G and Rob J Hyndman (2006). "25 years of time series forecasting". In: *International Journal of Forecasting* 22.3, pp. 443–473.

De Santis, Roberto A (2014). "The euro area sovereign debt crisis: Identifying flight-to-liquidity and the spillover mechanisms". In: *Journal of Empirical Finance* 26, pp. 150–170.

Dhanani, Alpa, Suzanne Fifield, Christine Helliar, and Lorna Stevenson (2007). "Why UK companies hedge interest rate risk". In: *Studies in Economics and Finance* 24.1, pp. 72–90.

Dick, Christian D, Ronald MacDonald, and Lukas Menkhoff (2015). "Exchange rate forecasts and expected fundamentals". In: *Journal of International Money and Finance* 53, pp. 235–256.

Diebold, Francis X. and Jose A. Lopez (1996). "Forecast evaluation and combination". In: *Statistical Methods in Finance*. Vol. 14. Handbook of Statistics. Elsevier, pp. 241–268.

Diebold, Francis X and Roberto S Mariano (1995). "Comparing predictive accuracy". In: *Journal of Business & Economic Statistics* 13.3, pp. 253–263.

Dixon, Robert, Zhichao Zhang, and Yang Dai (2016). "Exchange rate flexibility in China: Measurement, regime shifts and driving forces of change". In: *Review of International Economics* 24.5, pp. 875–892.

Dominguez, Kathryn M (1986). "Are foreign exchange forecasts rational?: New evidence from survey data". In: *Economics Letters* 21.3, pp. 277–281.

Dovern, Jonas (2015). "A multivariate analysis of forecast disagreement: Confronting models of disagreement with survey data". In: *European Economic Review* 80, pp. 16–35.

Dovern, Jonas and Johannes Weisser (2011). "Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7". In: *International Journal of Forecasting* 27.2, pp. 452–465.

Dovern, Jonas and Christina Ziegler (2008). "Predicting growth rates and recessions. Assessing US leading indicators under real-time conditions". In: *Applied Economics Quarterly* 54.4, pp. 293–318.

Dow, James and Sergio Ribeiro da Costa Werlang (1992). "Excess volatility of stock prices and Knightian uncertainty". In: *European Economic Review* 36.2-3, pp. 631–638.

Dreger, Christian and Georg Stadtmann (2008). "What drives heterogeneity in foreign exchange rate expectations: Insights from a new survey". In: *International Journal of Finance & Economics* 13.4, pp. 360–367.

Dua, Pami (1988). "Multiperiod forecasts of interest rates". In: *Journal of Business & Economic Statistics* 6.3, pp. 381–384.

Duffee, Gregory (2013). "Forecasting interest rates". In: *Handbook of Economic Forecasting* 2, pp. 385–426.
Düll, Robert, Felix König, and Jana Ohls (2017). "On the exposure of insurance companies to sovereign risk —- Portfolio investments and market forces". In: *Journal of Financial Stability* 31, pp. 93–106.
Elder, John and Apostolos Serletis (2010). "Oil price uncertainty". In: *Journal of Money, Credit and Banking* 42.6, pp. 1137–1159.
Elliott, Graham and Allan Timmermann (2008). "Economic forecasting". In: *Journal of Economic Literature*, pp. 3–56.
Engle, Robert F and Clive WJ Granger (1987). "Co-integration and error correction: Representation, estimation, and testing". In: *Econometrica: Journal of the Econometric Society*, pp. 251–276.
Epstein, Larry G and Tan Wang (1994). "Intertemporal asset pricing under Knightian uncertainty". In: *Econometrica: Journal of the Econometric Society*, pp. 283–322.
Estrella, Arturo and Frederic S Mishkin (1998). "Predicting U.S. recessions: Financial variables as leading indicators". In: *The Review of Economics and Statistics* 80.1, pp. 45–61.
Fama, Eugene F (1970). "Efficient capital markets: A review of theory and empirical work". In: *The Journal of Finance* 25.2, pp. 383–417.
Fatemi, Ali and Martin Glaum (2000). "Risk management practices of German firms". In: *Managerial Finance* 26.3, pp. 1–17.
Fauvel, Yvon, Alain Paquet, and Christian Zimmermann (1999). *A survey on interest rate forecasting*. Tech. rep. CREFE, Université du Québec à Montréal.
Feschiyan, Daniela and Radka Andasarova (2017). "Accounting analysis of the loan portfolio of banks and feasibility of the model for credit losses". In: *Economic Alternatives* 4, pp. 532–554.
Fildes, Robert, Konstantinos Nikolopoulos, Sven F Crone, and AA Syntetos (2008). "Forecasting and operational research: a review". In: *Journal of the Operational Research Society* 59.9, pp. 1150–1172.
Fintzen, David and HO Stekler (1999). "Why did forecasters fail to predict the 1990 recession?" In: *International Journal of Forecasting* 15.3, pp. 309–323.
Firth, Michael (1975). "The Role of forecasting in business". In: *Managerial Finance* 1.2, pp. 97–108.
Fisher, Irving (1907). *The rate of interest: Its nature, determination and relation to economic phenomena*. Macmillan.
Frankel, Jeffrey A and Shang-Jin Wei (2007). "Assessing China's exchange rate regime". In: *Economic Policy* 22.51, pp. 576–627.
Franses, Philip Hans, Michael McAleer, and Rianne Legerstee (2014). "Evaluating macroeconomic forecasts: A concise review of some recent developments". In: *Journal of Economic Surveys* 28.2, pp. 195–208.
Fratzscher, Marcel and Arnaud Mehl (2014). "China's dominance hypothesis and the emergence of a tri-polar global currency system". In: *The Economic Journal* 124.581, pp. 1343–1370.
Friedman, Benjamin M (1980). "Survey evidence on the rationality of interest rate expectations". In: *Journal of Monetary Economics* 6.4, pp. 453–465.
Friedman, Milton (1977). "Nobel lecture: inflation and unemployment". In: *Journal of Political Economy* 85.3, pp. 451–472.
Froot, Kenneth A, David S Scharfstein, and Jeremy C Stein (1993). "Risk management: Coordinating corporate investment and financing policies". In: *The Journal of Finance* 48.5, pp. 1629–1658.

Fujiwara, Ippei, Hibiki Ichiue, Yoshiyuki Nakazono, and Yosuke Shigemi (2013). "Financial markets forecasts revisited: Are they rational, stubborn or jumpy?" In: *Economics Letters* 118.3, pp. 526–530.

Funke, Michael, Chang Shu, Xiaoqiang Cheng, and Sercan Eraslan (2015). "Assessing the CNH–CNY pricing differential: Role of fundamentals, contagion and policy". In: *Journal of International Money and Finance* 59, pp. 245–262.

Gelper, Sarah and Christophe Croux (2007). "Multivariate out-of-sample tests for Granger causality". In: *Computational Statistics & Data Analysis* 51.7, pp. 3319–3329.

Giddy, Ian H and Gunter Dufey (1975). "The random behavior of flexible exchange rates: Implications for forecasting". In: *Journal of International Business Studies* 6.1, pp. 1–32.

Goldstein, Morris and Nicholas Lardy (2006). "China's exchange rate policy dilemma". In: *The American Economic Review* 96.2, pp. 422–426.

Goodwin, Paul and George Wright (2010). "The limits of forecasting methods in anticipating rare events". In: *Technological Forecasting and Social Change* 77.3, pp. 355–368.

Gramlich, Dieter and Mikhail V Oet (2011). "The structural fragility of financial systems: Analysis and modeling implications for early warning systems". In: *The Journal of Risk Finance* 12.4, pp. 270–290.

Granger, Clive WJ (1981). "Some properties of time series data and their use in econometric model specification". In: *Journal of Econometrics* 16.1, pp. 121–130.

Granger, Clive WJ and M Hashem Pesaran (2000). "Economic and statistical measures of forecast accuracy". In: *Journal of Forecasting* 19.7, pp. 537–537.

Granger, Clive WJ and Ramu Ramanathan (1984). "Improved methods of combining forecasts". In: *Journal of Forecasting* 3.2, pp. 197–204.

Greer, Mark (2003). "Directional accuracy tests of long-term interest rate forecasts". In: *International Journal of Forecasting* 19.2, pp. 291–298.

Greer, Mark R (2005). "Combination forecasting for directional accuracy: An application to survey interest rate forecasts". In: *Journal of Applied Statistics* 32.6, pp. 607–615.

Gripaios, Peter (1994). "The use and abuse of economic forecasts". In: *Management Decision* 32.6, pp. 61–64.

Gubaydullina, Zulia, Oliver Hein, and Markus Spiwoks (2011). "The status quo bias of bond market analysts". In: *Journal of Applied Finance and Banking* 1.1, p. 31.

Hartmann, Matthias, Helmut Herwartz, and Maren Ulm (2017). "A comparative assessment of alternative ex ante measures of inflation uncertainty". In: *International Journal of Forecasting* 33.1, pp. 76–89.

Harvey, Campbell R (1991). "Interest rate based forecasts of German economic growth". In: *Review of World Economics* 127.4, pp. 701–718.

Hendry, David F and Michael P Clements (2003). "Economic forecasting: Some lessons from recent research". In: *Economic Modelling* 20.2, pp. 301–329.

— (2004). "Pooling of forecasts". In: *The Econometrics Journal* 7.1, pp. 1–31.

Hernandez, Leonardo and Peter J Montiel (2003). "Post-crisis exchange rate policy in five Asian countries: Filling in the "hollow middle"". In: *Journal of the Japanese and International Economies* 17.3, pp. 336–369.

Holden, Ken, David A Peel, and John L Thompson (1990). *Economic forecasting: an introduction*. Cambridge University Press.

Howard, John A (1954). "A note on corporate forecasting practices". In: *The Journal of Business* 27.1, pp. 101–105.

Hu, May, Yunfeng Li, Jingjing Yang, and Chi-Chur Chao (2016). "Actual intervention and verbal intervention in the Chinese RMB exchange rate". In: *International Review of Economics & Finance* 43, pp. 499–508.

Ince, Onur and Tanya Molodtsova (2017). "Rationality and forecasting accuracy of exchange rate expectations: Evidence from survey-based forecasts". In: *Journal of International Financial Markets, Institutions and Money* 47, pp. 131–151.

Istrefi, K and S Mouabbi (2017). *Subjective interest rate uncertainty and the macroeconomy: A cross-country analysis*. Tech. rep. Banque de France.

Ito, Takatoshi (1990). "Foreign exchange rate expectations: Micro survey data". In: *The American Economic Review*, pp. 434–449.

— (2010). "China as number one: How about the Renminbi?" In: *Asian Economic Policy Review* 5.2, pp. 249–276.

Jo, Soojin (2014). "The effects of oil price uncertainty on global real economic activity". In: *Journal of Money, Credit and Banking* 46.6, pp. 1113–1135.

Jongen, Ron, Willem FC Verschoor, and Christian CP Wolff (2008). "Foreign exchange rate expectations: Survey and synthesis". In: *Journal of Economic Surveys* 22.1, pp. 140–165.

Jorion, Philippe (1990). "The exchange-rate exposure of U.S. multinationals". In: *Journal of Business*, pp. 331–345.

Jurado, Kyle, Sydney C Ludvigson, and Serena Ng (2015). "Measuring uncertainty". In: *American Economic Review* 105.3, pp. 1177–1216.

Kan, Yoke Yue (2017). "Why RMB should be more flexible". In: *Journal of Financial Economic Policy* 9.02, pp. 156–173.

Keynes, John Maynard (1936). *The General Theory of Employment, Interest and Money*. London: Macmillan.

Khan, Muhammad Imran (2017). "Falling oil prices: Causes, consequences and policy implications". In: *Journal of Petroleum Science and Engineering* 149, pp. 409–427.

Kilian, Lutz (2009). "Not all oil Price shocks are alike: Disentangling demand and supply shocks in the crude oil market". In: *The American Economic Review* 99.3, p. 1053.

Kim, Don H and Athanasios Orphanides (2012). "Term structure estimation with survey data on interest rate forecasts". In: *Journal of Financial and Quantitative Analysis* 47.1, pp. 241–272.

Klose, Jens and Benjamin Weigert (2014). "Sovereign yield spreads during the euro crisis: Fundamental factors versus redenomination risk". In: *International Finance* 17.1, pp. 25–50.

Knight, Frank H (1921). "Risk, uncertainty and profit". In: *New York: Hart, Schaffner and Marx*.

Kolb, RA and Harold O Stekler (1996). "How well do analysts forecast interest rates?" In: *Journal of Forecasting* 15.5, pp. 385–394.

Kunze, Frederik (2014). "The European financial Crisis – A Challenge for ten-year German government bond yield dorecasts?" In: *International Journal of Bonds and Derivatives* 1.2, pp. 171–185.

Kunze, Frederik and Mario Gruppe (2014). "Performance of survey forecasts by professional analysts: Did the European debt crisis make it harder or perhaps even easier?" In: *Social Sciences* 3.1, pp. 128–139.

Kunze, Frederik, Mario Gruppe, and Tilo Wendler (2015). "Interest rate forecasting and the financial crisis: A turning point in more than just one way". In: *International Journal of Financial Engineering and Risk Management* 2.1, pp. 1–16.

Kunze, Frederik, Jens Kramer, and Norman Rudschuck (2014). "Interest rate forecasts in times of financial crisis: What might be interesting to know?" In: *European Journal of Political Economy* 34, S45–S52.

Kuznets, Simon (1954). "Concepts and assumptions in long-term projections of national product". In: *Long-Range Economic Projection*. NBER, pp. 7–42.

Lahiri, Kajal and Xuguang Sheng (2010). "Measuring forecast uncertainty by disagreement: The missing link". In: *Journal of Applied Econometrics* 25.4, pp. 514–538.

Lee, Chi-Chuan and Chien-Chiang Lee (2018). "Oil price shocks and Chinese banking performance: Do country risks matter?" In: *Energy Economics*.

Liu, Peter C and Gangadharrao S Maddala (1992). "Rationality of survey data and tests for market efficiency in the foreign exchange markets". In: *Journal of International Money and Finance* 11.4, pp. 366–381.

Liu, Tao (2015). *The onshore-offshore interaction of RMB market: A high-frequency analysis*. MPRA Paper 64057. Munich Personal RePEc Archive.

Lo, Andrew W. (1997). "A non-random walk down Wall Street". In: *Proceedings of Symposia in Pure Mathematics*. Vol. 60, pp. 149–183.

Lopez-Martin, Bernabe, Julio Leal, and Andre Martinez Fritscher (2017). "Commodity price risk management and fiscal policy in a sovereign default model". In: *Journal of International Money and Finance*.

Lowe, EA (1970). "Efficient forecasting and management information systems". In: *Management Decision* 4.4, pp. 51–54.

Lubik, Thomas A and Frank Schorfheide (2007). "Do central banks respond to exchange rate movements? A structural investigation". In: *Journal of Monetary Economics* 54.4, pp. 1069–1087.

Lucarelli, Stefano, Filippo Umberto Andrini, and Annamaria Bianchi (2018). "Euro depreciation and trade asymmetries between Germany and Italy versus the U.S.: Industry-level estimates". In: *Applied Economics* 50.1, pp. 15–34.

Lucas Jr, Robert E (1976). "Econometric policy evaluation: A critique". In: *Carnegie-Rochester Conference Series on Public Policy*. Vol. 1. Elsevier, pp. 19–46.

Lütkepohl, Helmut and Markus Krätzig (2004). *Applied time series econometrics*. Cambridge University Press.

Ma, Guonan and Robert N McCauley (2008). "Efficacy of China's capital controls: Evidence from price and flow data". In: *Pacific Economic Review* 13.1, pp. 104–123.

MacDonald, Ronald (2000). "Expectations formation and risk in three financial markets: Surveying what the surveys say". In: *Journal of Economic Surveys* 14.1, pp. 69–100.

MacDonald, Ronald and Thomas S Torrance (1988). "Exchange rates and the "news": Some evidence from UK survey data". In: *The Manchester School* 56.1, pp. 69–76.

Madura, Jeff and Emilio R Zarruk (1995). "Bank exposure to interest rate risk: A global perspective". In: *Journal of Financial Research* 18.1, pp. 1–13.

Magkonis, Georgios and Andreas Tsopanakis (2016). "The financial and fiscal stress interconnectedness: The case of G5 economies". In: *International Review of Financial Analysis* 46, pp. 62–69.

Makridakis, Spyros (1986). "The art and science of forecasting An assessment and future directions". In: *International Journal of Forecasting* 2.1, pp. 15–39.

— (1996). "Forecasting: Its role and value for planning and strategy". In: *International Journal of Forecasting* 12.4, pp. 513–537.

Makridakis, Spyros, Michele Hibon, and Claus Moser (1979). "Accuracy of forecasting: An empirical investigation". In: *Journal of the Royal Statistical Society. Series A (General)*, pp. 97–145.

Mankiw, N Gregory, Ricardo Reis, and Justin Wolfers (2003). "Disagreement about inflation expectations". In: *NBER Macroeconomics Annual* 18, pp. 209–248.

McCandless Jr, George T and Warren E Weber (1995). "Some monetary facts". In: *Federal Reserve Bank of Minneapolis. Quarterly Review-Federal Reserve Bank of Minneapolis* 19.3, p. 2.

McNees, Stephen K (1988). "On the future of macroeconomic forecasting". In: *International Journal of Forecasting* 4.3, pp. 359–362.

Meub, Lukas, Till Proeger, and Kilian Bizer (2013). *Anchoring: A valid explanation for biased forecasts when rational predictions are easily accessible and well incentivized?* Tech. rep. Discussion Papers, Center for European Governance and Economic Development Research.

Miao, Hong, Sanjay Ramchander, Tianyang Wang, and Dongxiao Yang (2017). "Influential factors in crude oil price forecasting". In: *Energy Economics* 68, pp. 77–88.

Mills, Terence C (2008). "Forecasting financial variables". In: *A Companion to Economic Forecasting*. Ed. by Michael P Clements and David F Hendry. United Kingdom: Blackwell Publishing, pp. 510–538.

Mishkin, Frederic S (1995). "Symposium on the monetary transmission mechanism". In: *The Journal of Economic Perspectives* 9.4, pp. 3–10.

Mitchell, Karlyn and Douglas K Pearce (2007). "Professional forecasts of interest rates and exchange rates: Evidence from the Wall Street Journal's panel of economists". In: *Journal of Macroeconomics* 29.4, pp. 840–854.

Mohaddes, Kamiar and M Hashem Pesaran (2017). "Oil prices and the global economy: Is it different this time around?" In: *Energy Economics*.

Moore, Angus (2017). "Measuring economic uncertainty and its effects". In: *Economic Record* 93.303, pp. 550–575.

Morecroft, John DW (2017). "Enduring feedback structure and long-term futures for the upstream oil Industry". In: *Systems Research and Behavioral Science* 34.4, pp. 494–509.

Morrison, Wayne M (2011). "China-U.S. trade issues". In: *Current Politics and Economics of Northern and Western Asia* 20.3, p. 409.

Muth, John F (1961). "Rational expectations and the theory of price movements". In: *Econometrica: Journal of the Econometric Society*, pp. 315–335.

Neuenkirch, Matthias (2012). "Managing financial market expectations: The role of central bank transparency and central bank communication". In: *European Journal of Political Economy* 28.1, pp. 1–13.

Nordhaus, William D (1987). "Forecasting efficiency: Concepts and applications". In: *The Review of Economics and Statistics*, pp. 667–674.

Novotny-Farkas, Zoltán (2016). "The interaction of the IFRS9 expected loss approach with supervisory rules and implications for financial stability". In: *Accounting in Europe* 13.2, pp. 197–227.

Nusair, Salah A (2016). "The effects of oil price shocks on the economies of the gulf co-operation council countries: Nonlinear analysis". In: *Energy Policy* 91, pp. 256–267.

Ouyang, Alice Y and Ramkishen S Rajan (2011). "Reserve accumulation and monetary sterilization in Singapore and Taiwan". In: *Applied Economics* 43.16, pp. 2015–2031.

Pagan, Adrian R and John Robertson (2008). “Forecasting for policy”. In: *A Companion to Economic Forecasting*. Ed. by Michael P Clements and David F Hendry. United Kingdom: Blackwell Publishing, pp. 152–178.
Parsley, David C and Helen A Popper (2006). “Exchange rate pegs and foreign exchange exposure in East and South East Asia”. In: *Journal of International Money and Finance* 25.6, pp. 992–1009.
Perron, Pierre (1990). “Testing for a unit root in a time series with a changing mean”. In: *Journal of Business & Economic Statistics* 8.2, pp. 153–162.
— (2006). “Dealing with structural Breaks”. In: *Palgrave Handbook of Econometrics* 1.2, pp. 278–352.
Perron, Pierre and Timothy J Vogelsang (1992). “Testing for a unit root in a time series with a changing mean: corrections and extensions”. In: *Journal of Business & Economic Statistics* 10.4, pp. 467–470.
Pesaran, M Hashem and Allan Timmermann (2004). “How costly is it to ignore breaks when forecasting the direction of a time series?” In: *International Journal of Forecasting* 20.3, pp. 411–425.
Pesaran, M Hashem and Martin Weale (2006). “Survey expectations”. In: *Handbook of Economic Forecasting* 1. Ed. by G. Elliott, C.W.J. Granger, and A. Timmermann, pp. 715–776.
Pollard, Patricia S (2003). “A look inside two central banks: The European Central Bank and the Federal Reserve”. In: *Federal Reserve Bank of St. Louis Review* 85.January/February 2003.
Poncela, Pilar and Eva Senra (2017). “Measuring uncertainty and assessing its predictive power in the Euro area”. In: *Empirical Economics* 53.1, pp. 165–182.
Pontines, Victor and Ramkishen S Rajan (2011). “Foreign exchange market intervention and reserve accumulation in emerging Asia: Is there evidence of fear of appreciation?” In: *Economics Letters* 111.3, pp. 252–255.
Prasad, Eswar S (2009). “Is the Chinese growth miracle built to last?” In: *China Economic Review* 20.1, pp. 103–123.
Prat, Georges and Remzi Uctum (2011). “Modelling oil price expectations: Evidence from survey data”. In: *The Quarterly Review of Economics and Finance* 51.3, pp. 236–247.
Proeger, Till and Lukas Meub (2014). “Overconfidence as a social bias: Experimental evidence”. In: *Economics Letters* 122.2, pp. 203–207.
Rajan, Ramikishen (2002). “Exchange rate policy options for post-crisis Southeast Asia: Is there a case for currency baskets?” In: *The World Economy* 25.1, pp. 137–163.
Ramirez, Carlos D (2013). “The political economy of “currency manipulation” bashing”. In: *China Economic Review* 27, pp. 227–237.
Reeves, Rachel and Michael Sawicki (2007). “Do financial markets react to Bank of England communication?” In: *European Journal of Political Economy* 23.1, pp. 207–227.
Reitz, Stefan, Jan-Christoph Rülke, and Georg Stadtmann (2009). *Are oil price forecasters finally right? Regressive expectations toward more fundamental values of the oil price*. Working Paper Series in Economics 09/04. WHU – Otto Beisheim School of Management.
Rigotti, Luca and Chris Shannon (2005). “Uncertainty and risk in financial markets”. In: *Econometrica* 73.1, pp. 203–243.
Ronald, Macdonald and Marsh Ian (1993). “On the efficiency of oil price forecasts”. In: *Applied Financial Economics* 3.4, pp. 293–302.

Roos, Charles F (1955). "Survey of economic forecasting techniques: A survey article". In: *Econometric, Journal of the Econometric Society*, pp. 363–395.
Rossi, Barbara and Tatevik Sekhposyan (2016). "Forecast rationality tests in the presence of instabilities, with applications to Federal Reserve and survey forecasts". In: *Journal of Applied Econometrics* 31.3, pp. 507–532.
Ruelke, Jan Christoph, Christian Pierdzioch, and Georg Stadtmann (2012). "On the internal consistency of short-term, medium-term and long-term oil price forecasts". In: *Applied Economics* 44.21, pp. 2757–2765.
Schultz, Randall L (1984). "The implementation of forecasting models". In: *Journal of Forecasting* 3.1, pp. 43–55.
Schwartz, Robert A (1970). "Efficient Capital Markets: A review of theory and empirical work: Discussion". In: *The Journal of Finance* 25.2, pp. 421–423.
Schwarzbach, Christoph, Frederik Kunze, Norman Rudschuck, and Torsten Windels (2012). "Asset management in the German insurance industry: The quality of interest rate forecasts". In: *Zeitschrift für die gesamte Versicherungswissenschaft* 101.5, pp. 693–703.
Shaeri, Komeil, Cahit Adaoglu, and Salih T Katircioglu (2016). "Oil price risk exposure: A comparison of financial and non-financial subsectors". In: *Energy* 109, pp. 712–723.
Shahzad, Syed Jawad Hussain, Nader Naifar, Shawkat Hammoudeh, and David Roubaud (2017). "Directional predictability from oil market uncertainty to sovereign credit spreads of oil-exporting countries: Evidence from rolling windows and crossquantilogram analysis". In: *Energy Economics* 68, pp. 327–339.
Sibbertsen, Philipp, Christoph Wegener, and Tobias Basse (2014). "Testing for a break in the persistence in yield spreads of EMU government bonds". In: *Journal of Banking & Finance* 41, pp. 109–118.
Sill, Keith (2014). "Forecast disagreement in the Survey of Professional Forecasters". In: *Business Review* Q2, pp. 15–24.
Singleton, Kenneth J (2013). "Investor flows and the 2008 boom/bust in oil prices". In: *Management Science* 60.2, pp. 300–318.
Skoglund, Jimmy (2017). "Credit risk term-structures for lifetime impairment forecasting: A practical guide". In: *Journal of Risk Management in Financial Institutions* 10.2, pp. 177–195.
Spiwoks, Markus, Nils Bedke, and Oliver Hein (2008). "Forecasting the past: The case of U.Ss interest rate forecasts". In: *Financial Markets and Portfolio Management* 22.4, pp. 357–379.
— (2010). "Topically orientated trend adjustment and autocorrelation of the residuals – An empirical investigation of the forecasting behavior of bond market analysts in Germany". In: *Journal of Money, Investment and Banking* 14, pp. 16–35.
Spiwoks, Markus and Oliver Hein (2007). "Die Währungs-, Anleihen-und Aktienmarktprognosen des Zentrums für Europäische Wirtschaftsforschung". In: *AStA Wirtschafts-und Sozialstatistisches Archiv* 1.1, pp. 43–52.
Stock, James H and Mark W Watson (1996). "Evidence on structural instability in macroeconomic time series relations". In: *Journal of Business & Economic Statistics* 14.1, pp. 11–30.
Stockman, Alan C (1987). "Economic theory and exchange rate forecasts". In: *International Journal of Forecasting* 3.1, pp. 3–15.
Sturm, Jan-Egbert and Jakob De Haan (2011). "Does central bank communication really lead to better forecasts of policy decisions? New evidence based on a Taylor rule model for the ECB". In: *Review of World Economics* 147.1, pp. 41–58.

Subramanian, Arvind and Martin Kessler (2013). "The Renminbi bloc is here: Asia down, rest of the world to go?" In: *Journal of Globalization and Development* 4.1, pp. 49–94.

Taylor, John B (2001). "The role of the exchange rate in monetary-policy rules". In: *American Economic Review* 91.2, pp. 263–267.

Theil, Henri (1955). "Who forecasts best?" In: *International Economic Papers* 5, pp. 194–199.

— (1992). "Who forecasts best?" In: *Henri Theil's Contributions to Economics and Econometrics: Volume III: Economic Policy and Forecasts, and Management Science*. Ed. by Baldev Raj and Johan Koerts. Springer Netherlands, pp. 1115–1120.

Timmermann, Allan and Clive WJ Granger (2004). "Efficient market hypothesis and forecasting". In: *International Journal of Forecasting* 20.1, pp. 15–27.

Trabelsi, Emna (2016). "Central bank transparency and the consensus forecast: What does The Economist poll of forecasters tell us?" In: *Research in International Business and Finance* 38, pp. 338–359.

Tsai, Chun-Li (2015). "How do U.S. stock returns respond differently to oil price shocks pre-crisis, within the financial crisis, and post-crisis?" In: *Energy Economics* 50, pp. 47–62.

Van Der Ploeg, Frederick (2017). "Macro policy responses to natural resource windfalls and the crash in commodity prices". In: *Journal of International Money and Finance*.

Vogelsang, Timothy J and Pierre Perron (1998). "Additional tests for a unit root allowing for a break in the trend function at an unknown time". In: *International Economic Review*, pp. 1073–1100.

von Spreckelsen, Christian, Frederik Kunze, Torsten Windels, and Hans-Jörg von Mettenheim (2014). "Forecasting Renminbi quotes in the revised Chinese FX market – can we get implications for the onshore/offshore spread-behaviour?" In: *International Journal of Economic Policy in Emerging Economies* 7.1, pp. 66–76.

West, Kenneth D (2006). "Forecast evaluation". In: *Handbook of economic forecasting*. Ed. by G. Elliott, C.W.J. Granger, and A. Timmermann. Vol. 1. Elsevier, pp. 99–134.

Wheelock, David C and Mark E Wohar (2009). "Can the term spread predict output growth and recessions? A survey of the literature". In: *Federal Reserve Bank of St. Louis Review* 91.5 Part 1, pp. 419–440.

Wheelwright, Steven C and Darral G Clarke (1976). "Corporate forecasting: Promise and reality." In: *Harvard Business Review* 54.6, pp. 40–64.

Winklhofer, Heidi, Adamantios Diamantopoulos, and Stephen F Witt (1996). "Forecasting practice: A review of the empirical literature and an agenda for future research". In: *International Journal of Forecasting* 12.2, pp. 193–221.

Zarnowitz, Victor and Louis A Lambros (1987). "Consensus and uncertainty in economic prediction". In: *Journal of Political Economy* 95.3, pp. 591–621.

Zhang, Ming and Bin Zhang (2017). "The boom and bust of the RMB's internationalization: A perspective from cross-border arbitrage". In: *Asian Economic Policy Review* 12.2, pp. 237–253.

Chapter 2

Asset liability management and the euro crisis -– sovereign credit risk as a challenge for the German life insurance industry

with Miguel Rodriguez, Christoph Schwarzbach and Christoph Dieng
contribution Frederik Kunze: 60%

Published:
The Journal of Risk Finance, Vol. 18 Issue: 4, pp. 466-483,
https://doi.org/10.1108/JRF-01-2017-0016

Abstract

Purpose — This paper aims to investigate the long-term relationships of long-term European Monetary Union (EMU) government bond yields. From an asset managers' or risk managers' perspective during the euro crisis, the relevance of sovereign credit and redenomination risk became a major issue. Furthermore, it has to be differentiated between core and non-core EMU member countries.

Design/methodology/approach -– Methods of applied time series analysis are used to investigate EMU government bond yields and EMU government bond yield spreads for Spain, Italy, The Netherlands, Austria and Germany. Both standard unit root testing procedures and breakpoint unit root tests are used to examine cointegrating relationships and structural changes in these relationships.

Findings -– The empirical results deliver clear evidence for structural shifts in the long-term relationship between Germany and the two non-core EMU countries (Italy and Spain). The timing of the breaks coincides with the timing of the euro crisis. On the contrary, the results for Austria and The Netherlands are different from the findings for the two non-core countries.

Research limitations/implications –- One major limitation of the study is the limited availability of data regarding to the reaction of asset managers or risk managers to the euro crisis. Especially in the context of the discussion with regard to the relevant risk-free rate for investors, this strand of research is relatively new.

Practical implications –- A deeper understanding of changes in the long-term relationship between government bond yields and the re-emergence of redenomination risk is important for asset managers and risk managers in the financial services industry. This is especially true for German life insurers.

Originality/value –- The study provides various empirical contributions to the literature on the euro crisis and sovereign credit risk. First, previous results with regard to the structural changes in the long-term relationship between German and Spanish, German and Italian, German and Austrian as well as Germany and Dutch government bond yields are confirmed using unit root breakpoint tests. Second, investigating the autoregressive coefficient and the timing of the breaks delivers evidence that non-core countries have been more exposed to the fear of redenomination risk. Third, we raise the question which risk free interest rate is relevant for the affected countries.

Keywords

Sovereign credit risk, asset liability management, European monetary union, life insurance sector, redenomination risk

Highlights

- We discuss the consequences of the EMU sovereign debt crisis for decision-makers in the financial services industry focusing on German life insurers.
- We apply unit root breakpoint tests to detect crisis related structural changes in the yields spreads between core and non-core EMU countries.
- The timings of the detected breaks deliver evidence that non-core countries seem to be more exposed to credit risk and redenomination risk.
- The results of our approach correspond to earlier findings in the literature and are of special interest for asset managers in the German life insurance industry investing in EMU non-core bonds.

1. Introduction

In the course of the global financial crisis and the following euro crisis starting in 2007 and 2010, respectively, asset and risk managers in the German financial services industry had to deal with demanding challenges. The upcoming fears of sovereign defaults, the perceived threat of a breakup of the European Monetary Union (EMU) as well as the comeback of redenomination risks placed asset values under severe pressure. This has especially affected German life insurers because they strongly invest in EMU sovereign bonds and related assets.

Life insurance policies hold strong significance for structuring the provision for old-age and death for the general population in Germany. Hence, they fulfill an important task for the German economy as a whole. For example, in 2015 there was a total of 91 million life insurance policies in Germany, which has a total population of around 82 million (GDV, 2016). Due to the large share of interest-bearing assets, the success of portfolio managers in the life insurance industry strongly depends on the interest rate level. Furthermore, the guaranteed interest rates included in many insurance contracts have to be generated on capital markets in the long term and as risk-free as possible. Unsurprisingly, EMU government bonds are an important asset class for German life insurance companies. Following official data from the German Federal Financial Supervisory Authority (Bundesanstalt fuer Finanzdienstleistungsaufsicht - BaFin), German life insurers had EUR 94 billion invested in EMU assets excluding Germany in Q4/2015.

The euro crisis had a huge influence on the assets of German life insurers. Given the significant role of EMU government bonds for insurance companies in Germany, a deeper understanding of the possible driving forces of EMU government bond yields remains to be very important and makes some further considerations inevitable. Structural shifts in the long-term relationship of EMU government bond yields have consequences for the value of all respective assets. This is also the case for German government bond yields, which are at a record low due to safe haven effects, deflationary fears caused by the crisis and the current monetary policy in the euro zone. Hence, relying solely on German government bonds would have severe implications for life insurers' interest income. However, compared to German government bonds, the higher yielding assets issued by Spain and Italy come with additional risks. In this context, not only the pricing of sovereign credit risk (SCR) – and hence fundamental factors like rising levels of public debt – but also the influence of redenomination risk has to be considered by asset and financial risk managers. In the case of a breakup of the euro zone, the liabilities of institutional investors (e.g. insurance companies in Germany) would be denominated in the new currency of the peripheral country that exits the EMU. Therefore, holding the redenominated government bonds would create losses due to currency devaluation.

As described above, risk and asset managers in the financial service industry and especially in the German life insurance industry have to deal with a challenging environment. Accordingly, we empirically examine developments in the EMU government bond market focusing on sovereign credit risk and redenomination risk using yield spreads between German and Spanish, Italian, Austrian and Dutch government bonds. Building upon existing studies mentioned later to reinvestigate the characteristics of EMU government bond yield spreads, we employ a different empirical approach testing for unit roots with breakpoints. We investigate the sovereign bond yields of the aforementioned countries to differentiate between so-called core and non-core (i.e. periphery) EMU countries. Artis and Zhang (2001) implemented this standard segmentation in the literature at an early stage, where Italy and Spain belong to the non-core or periphery countries whereas Germany, Austria and The Netherlands are seen as core countries (Antonakakis and Vergos, 2013; Angelopoulou et al., 2014; Claeys and Vašíček, 2014; Costantini et al. 2014).

The remainder of the paper is structured as follows. In Chapter 2, we briefly describe the asset allocation process of the German insurance industry with a focus on the current low interest rate environment, which leads directly to Chapter 3, providing an overview of the related literature and the European debt crisis. In Chapter 4, we present the data as well as the methodological framework. Chapter 5 presents and briefly discusses the empirical results. Finally, Chapter 6 concludes the paper and provides some ideas for further research.

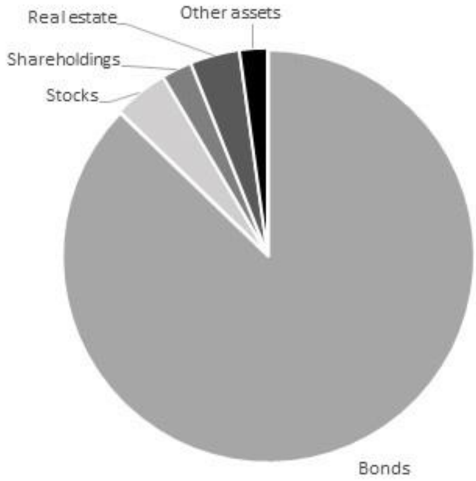
2. Asset allocation process in the German life insurance industry

To better understand the consequences of a low interest rate environment as well as the euro crisis on the investment patterns of life insurance companies, we first explain the fundamentals of their investment strategy. German life insurers differ from most other investors due to a long-term investment horizon and the necessity to achieve a return that covers long-term interest rate guarantees given to the clients. Their liabilities commonly have a duration of several decades. From an insurer's perspective, this also results in the need to invest the insurance premiums in the long term (Basse and Friedrich, 2008). However, a long-term investment period leads to corresponding challenges in the investment policy, particularly in connection with the need to comply with these guarantees. Potential risks mainly arise from changes in the interest rate and spreads (Bafin, 2010), which subsequently lead to (re-)investment risks. The mitigation of these source of risks is a major task for the life insurers' risk managers and their asset liability management (ALM). ALM's task is an integrated consideration and optimization of assets and liabilities, which has gained importance especially against the background of the regulatory standards of Solvency II (Schwarzbach et al., 2014).

ALM can comprise various techniques taking into account the sector-specific structural differences in the investment behavior of insurers. For life insurers, especially immunization strategies hold particular importance. One result of this is that life insurers primarily invest in low-risk, fixed interest rate investments. By using immunization strategies, mostly the risk of interest rate changes but also the other risks associated with these investments (liquidity risks a.s.o.) are supposed to be accounted for. One technique used by insurers is cashflow matching. The main idea of this procedure is a systematic comparison and matching of the cashflows expected to be generated from assets and required by liabilities. Put differently, the technique is used to ensure that the monetary inflows are able to cover the necessary outflows at any given time, which would alleviate or even extinguish the consequences of changes in, for example, the interest rate level. Another technique used by life insurers to minimize risk is duration matching. Thereby, efforts are made to equal out the interest rate sensitivities derived from the maturities of the payouts of the assets and liabilities. Accordingly, assets and liabilities will react equally to an interest rate change and thus will not affect the equity position of the life insurer, at least in the absence of other risks, e.g. exchange rate risk.

As previously mentioned, life insurers have to follow a particularly careful investment strategy due to long maturities in their liabilities and their products' important role for old-age provision. This behavior is also reflected in their investment allocation among the different types of assets. The overall balance sheet value of the German primary insurers amounted to a total of EUR 885 billion in 2015. Figure 1 shows the average capital stock of these insurance companies on December 31, 2015. The chart shows that the majority of investments (87.3%) are invested in low-risk possibly longer-term bonds. A significantly smaller proportion is invested in stocks (4.3%), longer term shareholdings (2.4%), real estate (3.9%) and other forms of investment (2.2%). The bonds can be further divided into a total of 25.9% being invested in bonds through funds, 20.5% in loans and 17.9% in covered bonds. This detailed account of their assets shows the low-risk and likely long-term investment behavior of life insurance companies in Germany.

Fig. 1. Average asset allocation of German life insurers (in %)



Own representation (based on GDV, 2016)

With this behavior in mind, one could expect a shift in the investments of German life insurers in the course of the recent crises to reduce risk. This reallocation could happen with a focus on either assets with a high quality (flight to quality) or the country of origin (flight home). For German insurers, the latter would imply an even stricter focus. In times of crisis, this flight home could also be interpreted as a way to increase the exposure to one's home and thereby raise the chances of being bailed out if necessary (Giannetti and Laeven, 2012). The shift to higher-quality assets by investors would mean pro-cyclical behavior, which is not observed for insurers in studies covering non-crises periods (Grinblatt and Keloharju, 2000; De Haan and Kakes, 2010; Ferreira and Matos, 2008). Nevertheless, as described in further detail in the following section, Bijlsma and Vermeulen (2016) found evidence for these tendencies by insurers from The Netherlands during the European sovereign debt crisis.

3. Literature overview and history of the euro crisis

In the past, SCR was mainly a problem for less-developed countries, although recently it has also become an issue for EMU member countries (Aizenman, 2013; Moro, 2014). Meanwhile, oil-producing countries also have to cope with this problem, as stated by Wegener et al., (2016). For example, Ang and Longstaff (2013) examined in detail the factors influencing SCR. In the context of analyzing the pricing of credit default swaps, they determine that SCR is strongly correlated with movements in financial markets. In addition, Afonso et al. (2012) use an event study to conclude that government bond yields react significantly to changes in rating grades. Furthermore, rating agencies' assessments are anticipated by market participants. All of this leads to a corresponding pricing of government bonds. These relationships have been crucially important for the EMU government bond market.

Since the start of the Euro in the 1990s, EMU government bond yields have been converging, which can be explained by an alignment of risks. The establishment of the EMU led to an elimination of exchange rate risks. Obviously, national risks were considered as equivalent. Accordingly, from 1999 to 2008 the government bond yields of the EMU member countries acted in an almost identical manner. For example, Basse et al. (2014) have shown empirically that a long-term relationship exists between German and Italian government bond yields. However, structural breaks have to be considered in this convergent development. These structural breaks occurred during the global financial crisis and the European debt crisis. The causes for the divergence of the yields can be explained by using the example of Spain. In the course of the global financial crisis in 2008 and the European debt crisis in 2012, the Spanish economy moved into a strong recession. This was reflected in heavily-rising unemployment rates and a negative trend in real gross domestic product. As a result, the long-term interest rates of Spanish government bonds sharply rose as the financial markets identified an increased default risk. The restructuring of Greek debt in March 2012 – through which creditors waived a large proportion of their nominal claims – played a crucial role for this realization.

Not only the increase in Spanish interest rates is remarkable, but also the simultaneous reduction in the long-term interest rates of German government bonds. This indicates that the default risks of German bonds were considered to be much lower than the Spanish ones. Therefore, investors assessed the risks of the EMU members as being increasingly heterogeneous and accordingly the pricing occurred without reference to each other. Furthermore, Sgherri and Zoli (2009) show that the assumption of banking risks by European countries implied hazardous linkages between financial market uncertainties and the evaluation of governmental debt default risk. Subsequently, the correlation between the Credit Default Swaps (CDS) of states and those of domestic banks increased. Since European countries issued guarantees for their domestic banks – which were subject to divergent default risks – the risk transfer to the sovereign default risks was also unequal and thus further increased the differences in the pricing of European government bonds.

Additionally, the perception of diverging risks and the respective pricing caused safe haven flows, which were problematic for the Southern European countries as they were particularly affected by the crisis. As a result of the apparently increased risk of Southern European bonds, other more risk-free long-term bonds were more frequently sought. Bernoth and Erdogan (2012) show that the majority of German and foreign investors found this ”safe haven” in Germany – Europe’s largest economy – which previously did not hold this status. Therefore, as explained above, the structural break in the convergence of the interest rates in 2011 to 2012 is not only justified by a rising risk of default of Southern European countries, but also by a reduction in German interest rates due to a ”flight to quality”, meaning that investors shift their capital away from riskier to safer (or as safe as possible) investments.

Some evidence of this behavior among insurance companies in this environment is also observed by Bijlsma and Vermeulen (2016), who examine over 60 Dutch insurance companies. Contrary to the results by other authors for earlier time periods, they find a pro-cyclical investment behavior among insurance companies during the crises: Southern European assets (especially government bonds) were reduced in favor of potentially more secure Northern European assets. This behavior was only found in periods of crises and diminished after European Central Bank (ECB) president Draghi’s historical ”whatever it takes” speech. This famous speech was part of a much wider effort by the ECB to counteract these interest rate developments. Even unconventional measures are seen as necessary because most standard procedures – such as lowering the key interest rate – have already been extensively utilized. One of these interventions in particular is quantitative easing, which has been carried out by the ECB since March 2015. Thereby, the ECB (mainly through the countries’ central banks) purchases government bonds worth EUR 60 billion each month. In doing so, government bonds from most EMU countries are bought up according to a capital key. One goal of this is to prevent deflation. However, at the same time, it assists the countries with their refinancing simply because the increased demand lowers the interest rates of government bonds. This especially helps the Southern European countries, although it also strengthens the German bonds, which already exhibit very low levels of return. The quantitative easing by the ECB causes the yields to decline even further. This shows a fundamental weakness of the ECB’s policy: it is only applied uniformly to all countries, but not individually to selected countries. While many countries with refinancing problems benefit from the ECB’s measures, it creates numerous problems for the German economy (i.e. ultra-low interest rates) and here especially for the risk and asset managers in German life insurance companies.

As a further consequence of the EMU currency crisis, the determinants of sovereign bond yields of EMU member countries gained strong attention among academic scholars and at the peak of the crisis the irreversibility of the Euro was questioned (De Haan et al, 2014; Chang and Leblond, 2015). One important implication of the euro crisis has been the necessity to distinguish between fundamental factors and redenomination risk (Klose and Weigert, 2014) when dealing with sovereign yield spreads. Given the economic relevance of this topic for the financial service industry and academic researchers, there already exists a strand of literature discussing the implications of redenomination risk (Cœuré, 2013; Sibbertsen et al., 2014; De Backer, 2015; Favero and Missale, 2016). Even prior to the outbreak of the EMU crisis and the preceding financial shock in 2008/09, the debt levels rapidly increased in numerous euro countries (Lane, 2012). In fact, following the economic theory a rising debt position should lead to higher risk premiums and hence should result in higher government bond yields (Bernoth et al., 2012).

However, redenomination risk and fears regarding a breakup of the EMU have obviously long been put aside by financial market participants (i.e. both asset and financial risk managers). In the early years, the EMU market participants did not differentiate between countries as issuers of debt (Bernoth and Erdogan, 2012). However, the default of Lehman Brothers and the EMU sovereign debt crisis definitely have to be seen as an inflection point (Maltritz, 2012; Basse, 2014; Iara and Wolff, 2014). Furthermore, contagion effects have to be considered (Kilponen et al. 2012; Blatt et al., 2015) and a distinction between the different countries within the EMU is necessary (Von Hagen et al., 2011; Buechel, 2013; Afonso et al., 2014).

The findings of Basse et al. (2012) hold strong importance in this context since the authors show that government bond yields for Germany and Italy are cointegrated. However, the authors detect structural breaks in the long-term relationship between the two bond yields in early 2009. Gruppe and Lange (2014) investigate the long-term relationship between government bond yields of Germany and Spain and find evidence of a cointegrating relationship up to early 2009. In addition, Basse (2014) analyze inter alia government bond yields for Germany, Austria and The Netherlands and find empirical evidence for long-term relationships between German and Austrian as well as German and Dutch long-term government bond yields. However, especially in the case of the Netherlands, no structural break in the cointegrating relationship is found. Following Basse (2014), the timing of the break date in Austria in 2006 does not allow the conclusion that it is related to the sovereign debt crisis.

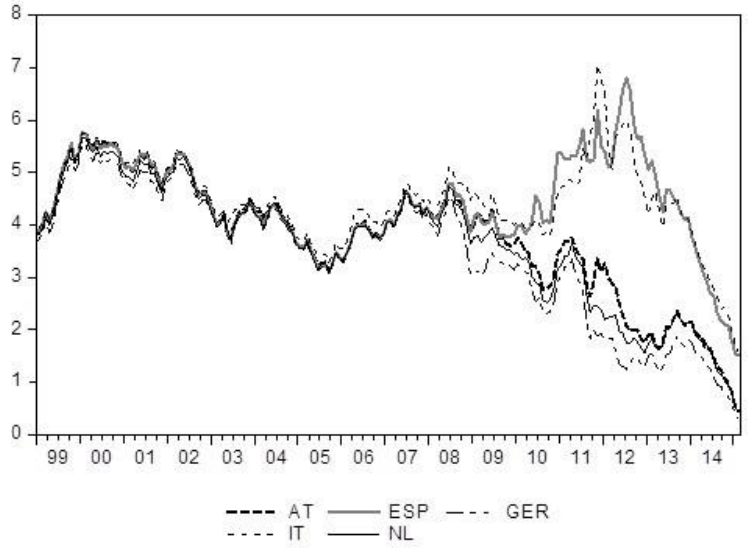
These studies hold strong relevance for the purpose of this paper. We also investigate the relationship between German and Spanish, German and Italian, German and Austrian as well as German and Dutch government bond yields, although we will use a different approach and focus on the corresponding bond yield spreads. Using the government bond yields of Germany as a benchmark is quite common in the literature due to the economic relevance of Germany as the largest EMU economy and its safe haven status (Canarella et al., 2011; Basse, 2014; MacDonald et al., 2015).

4. Data and methodology

For the empirical investigation, we use ten-year government bond yields from Germany as a reference country, from Italy and Spain as non-core as well as from Austria and the Netherlands as core EMU countries. The monthly data set was collected at Federal Reserve Economic Data (FRED), provided by the Federal Reserve Bank of St. Louis, covering the period from January 1999 to February 2015. Hence, it incorporates both the beginning of the EMU as well as the subprime and euro crisis.

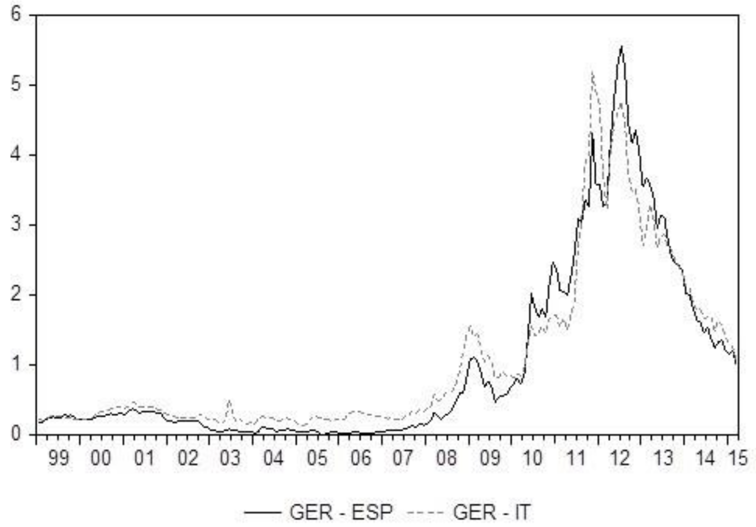
Since a long-term relationship between two time series can be empirically verified with the existence of a cointegrating relationship, we first check for unit roots in the bond yield time series. As Figure 2 shows, all interest rates seem to be non-stationary and hence they should share the same order of integration. To empirically test for unit roots, we apply the non-parametric testing procedure for unit roots proposed by Phillips and Perron (1988). Subsequently, we investigate the sovereign bond yield spreads (i.e. the simple differences between the corresponding two time series) for Germany and Spain (GER - ESP), Germany and Italy (GER - IT), Germany and Austria (GER - AT) as well as Germany and the Netherlands (GER - NL). For an existing long-run relationship, the corresponding spread time series should be integrated of order zero I(0) and hence should be stationary. As Figures 3 (GER - ESP / GER - IT) and 4 (GER - AT / GER - NL) show, there seem to be structural shifts. Accordingly, using ordinary unit root tests like the Phillips Perron (PP)-test used for the yield levels, would lead to biased results. Thus, we will apply a unit root test allowing for structural shifts with unknown timing (Perron, 1990; Perron and Vogelsang, 1992; Vogelsang and Perron, 1998; Perron, 2006). We apply the test for a unit root (null hypothesis) against the alternative hypothesis of a stationary process using the model specification for innovational outliers (IO) and additive outliers (AO). In general, the IO model specifications are more appropriate for gradual changes over time, whereas the AO model specifications fit better for sudden changes (Perron, 1994; Harvie and Pahlavani, 2006). Given the purpose of our investigations and as a robustness check, we apply both specifications.

Fig. 2. Ten-year government bond yields Austria (AT), the Netherlands (NL), Germany (GER), Spain (ESP) and Italy (IT)



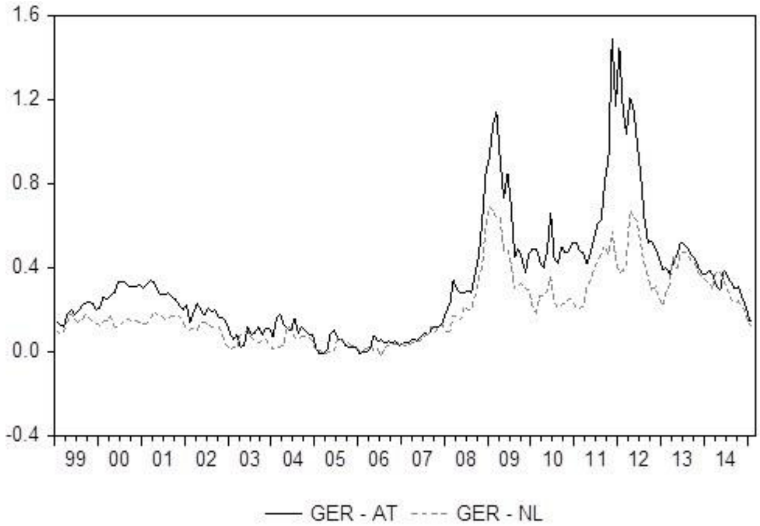
Own representation
based on FRED (2016)

Fig. 3. Spreads of ten-year government bond yields Spain vs. Germany (GER - ESP) and Italy vs. Germany (GER - IT)



Own representation
based on FRED (2016)

Fig. 4. Spreads of ten-year government bond yields Austria vs. Germany (GER - AT) and the Netherlands vs. Germany (GER - NL)



Own representation
based on FRED (2016)

5. Empirical evidence

The results of the PP-unit root tests for the EMU government bond yields for Germany, Italy, Spain, Austria and the Netherlands are reported in Tables 1 and 2 below. Following the empirical evidence, we can state that all four time series are non-stationary and integrated of order one or I(1), respectively. Given the empirical findings mentioned in the literature section, these results are unsurprising.

Table 1: PP-unit root tests in levels

		Critical value		
Country	Test stat.	1%	5%	10%
Austria	0.432645	-3.464280	-2.876356	-2.574746
Netherlands	0.361843	-3.464280	-2.876356	-2.574746
Spain	-0.329608	-3.464101	-2.876277	-2.574704
Italy	-0.386912	-3.464101	-2.876277	-2.574704
Germany	0.217802	-3.464101	-2.876277	-2.574704

Own
calculations

Table 2: PP-unit root tests in first differences Δ

		Critical value		
Country	Test stat.	1%	5%	10%
Austria	-10.66625	-3.464460	-2.876435	-2.574788
Netherlands	-10.43315	-3.464460	-2.876435	-2.574788
Spain	-11.24233	-3.464280	-2.876356	-2.574746
Italy	-10.90881	-3.464280	-2.876356	-2.574746
Germany	-10.24184	-3.464280	-2.876356	-2.574746

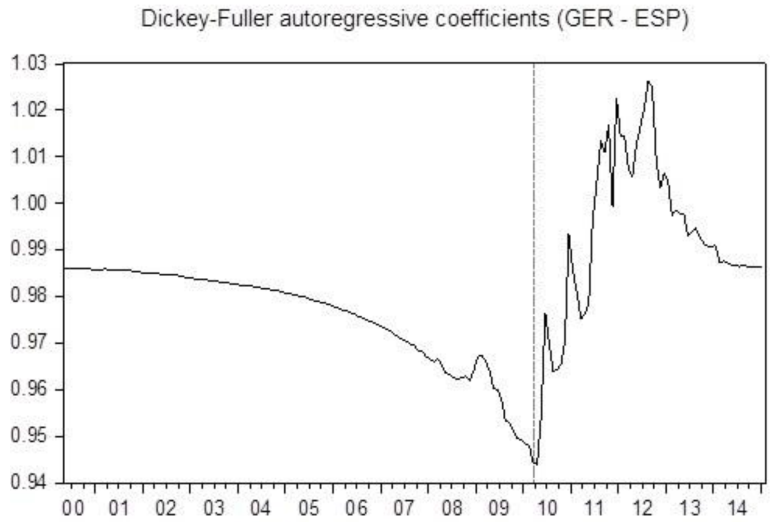
Own calculations

Figures 5 to 12 visualize the results of the breakpoint unit root tests for the spread time series. The figures show recursively estimated autoregressive processes. In the charts a coefficient value above one indicates a structural change in the cointegration relationship (i.e. spread time series as a linear combination of the corresponding government bond yields): in other words, an increase of the autoregressive coefficient above the benchmark implies that the government bond yield spreads for Italy and Spain do not remain stationary over time. In fact, the processes change from I(0) to I(1) in the course of the euro crisis. These results are also robust regarding the model specifications (i.e. IO or AO). Hence, for both non-core EMU countries, earlier findings of breaks in the long-term relationship have been empirically proven. These results are confirmed by the results shown in Tables 3 to 6 in the Appendix. For the non-core countries, we observe a breakpoint corresponding to the occurrence of the euro crisis. Again, these results are robust regarding the applied model specifications. For both the additive and innovation outlier, the break dates for Spain (IO: March 2010; AO: April 2011) and Italy (IO: March 2010; AO: February 2010) correspond to the timing of the euro crisis. We interpret the increasing autoregressive coefficient above the threshold value as an indication for the comeback of risk factors like redenomination risk (see Figures 5 and 6 for Spain and Figures 7 and 8 for Italy, respectively).

Regarding the EMU core countries, the results are somewhat different. In the case of the Netherlands, the null hypothesis of non-stationarity of the bond yield spread is rejected. More important in the context of our research are the break dates for Austria (IO: August 2008; AO: November 2007) and The Netherlands (IO: February 2008; AO: October 2008). Given the time of the euro crisis, these break dates may not be seen as a consequence of the euro crisis. Hence, for core EMU countries it seems to be the case that redenomination risk was not a major issue.

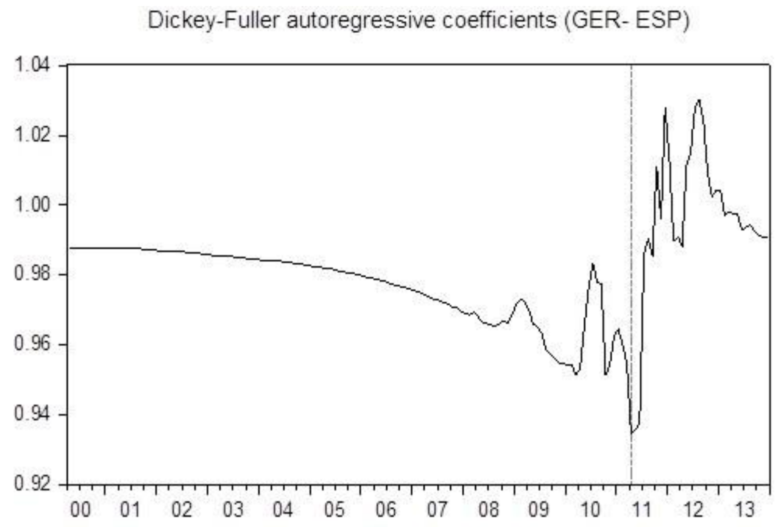
This interpretation gains further support from the estimated autoregressive coefficients shown in Figures 9 and 10 for Austria as well as Figures 11 and 12 for The Netherlands. However, in the case of Austria we observe a substantial increase in the autoregressive coefficient in the case of the innovation outlier specification. The coefficient reaches a value of 1. In fact the timing of this event does not seem to correspond to the euro crisis but to the financial turmoil in Austria attributed to the HETA case (Randl and Zechner, 2016). For The Netherlands, the occurrence of the break coincides with the ING Group facing financial difficulties due to the financial crisis, which led to government assistance for the group amounting to EUR 90 billion.

Fig. 5. Autoregressive coefficient; unit root break test GER - ESP; innovation outlier



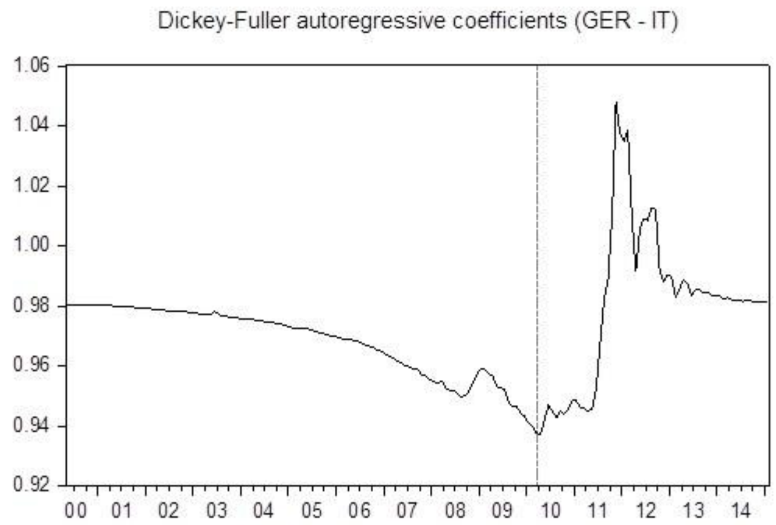
Own calculations

Fig. 6. Autoregressive coefficient; unit root break test GER - ESP; additive outlier



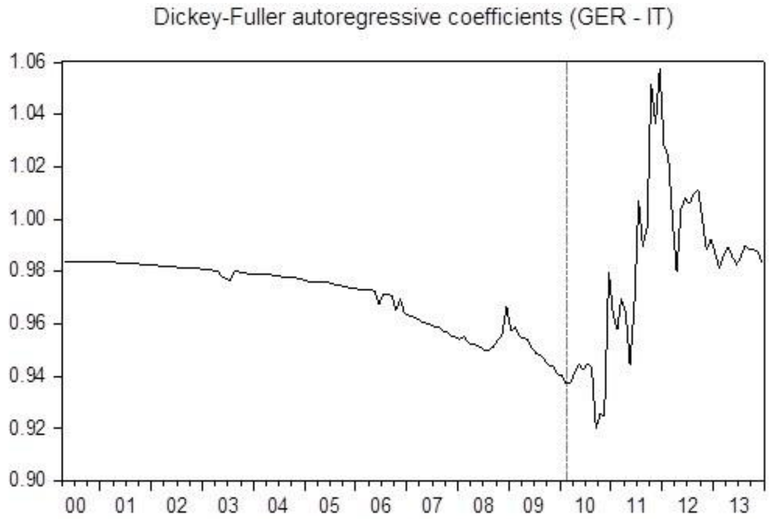
Own
calculations

Fig. 7. Autoregressive coefficient; unit root break test GER - IT; innovation outlier



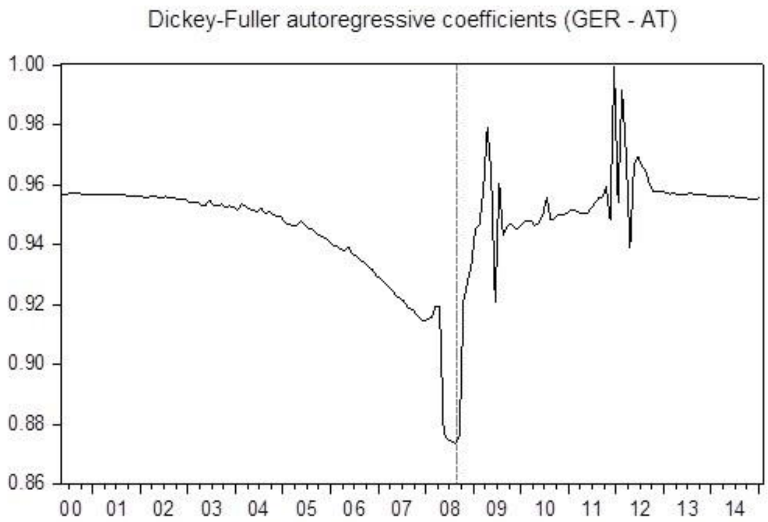
Own
calculations

Fig. 8. Autoregressive coefficient; unit root break test GER - IT; additive outlier



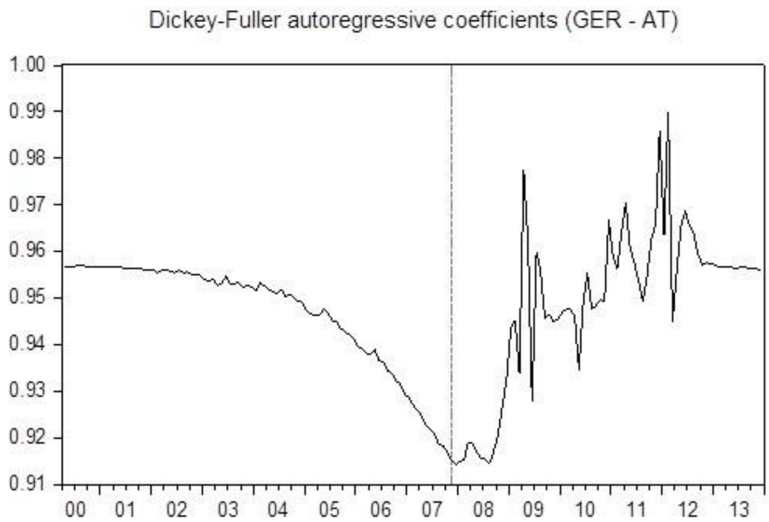
Own calculations

Fig. 9. Autoregressive coefficient; unit root break test GER - AT; innovation outlier



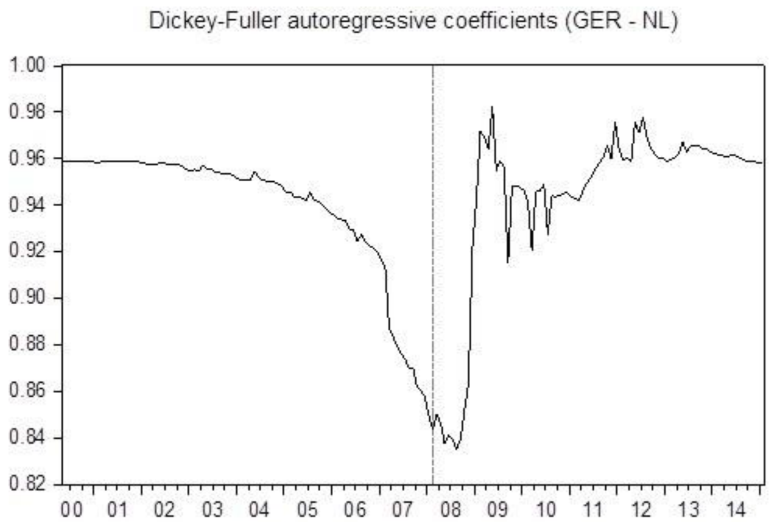
Own calculations

Fig. 10. Autoregressive coefficient; unit root break test GER - AT; additive outlier



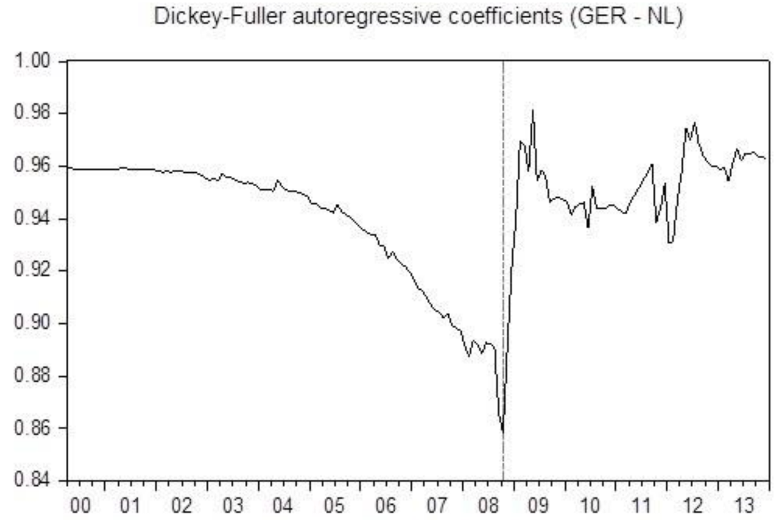
Own calculations

Fig. 11. Autoregressive coefficient; unit root break test GER - NL; innovation outlier



Own calculations

Fig. 12. Autoregressive coefficient; unit root break test GER - NL; additive outlier



Own calculations

6. Conclusions

In this paper, we have discussed the relevance of the EMU government bond market as well as EMU government bond yields for the financial service industry. One focus was placed on the ALM process of German life insurers. Furthermore, we have empirically investigated EMU government bond yield spreads. We assessed the pricing differential between German ten-year government bonds and two EMU core countries (The Netherlands and Austria) and two EMU non-core countries (Spain and Italy), respectively regarding structural changes using monthly data ranging from January 1999 to February 2015. We applied model specifications for innovation and additive outliers.

The statistical evidence from the utilized unit root breakpoint tests confirms previous findings regarding euro crisis-related structural shifts in the cointegrating relationship between German and Spanish and Italian government bond yields, respectively. Hence, based on our statistical results we have been able to show that sovereign credit and probably redenomination risks have led to structural changes in the relationship between government bond yields in peripheral countries and Germany. On the other hand, regarding the EMU core countries (in our case, The Netherlands and Austria), we have found no similar evidence of structural changes in the long-term relationship. First, the timing of the breakpoints for the yield spreads do not correspond with the euro crisis. Second – and even more importantly – in the context of our investigation, we have found no clear evidence for a structural change in the persistence of the bond yield spreads.

For risk and asset managers in the life insurance industry, these findings are important for various reasons. In fact, it has to be stated that especially redenomination risk is a challenge for asset managers especially in the German life insurance industry. In addition, further research should address the question of whether government bond yields can be seen as the risk-free rate for insurers in non-core EMU countries (in the context of our paper with respect to Spain and Italy) and core EMU countries (here: Germany, Austria and the Netherlands), respectively. In addition, the results also hold strong relevance for the financial services industry as a whole, since not only life insurers are heavily exposed to EMU countries. Furthermore, analyzing past structural breaks is helpful and necessary to develop a more profound understanding of matters in the context of EMU sovereign credit risk. Further research should focus on identifying potential upcoming structural shifts due to external or internal factors. Given the current alignment of the ECB's monetary policy, a more or less abrupt interest rate reversal may not be ruled out completely.

References

Afonso, A., Arghyrou, M. G. and Kontonikas, A., 2014. Pricing Sovereign Bond Risk in the European Monetary Union Area: An empirical Investigation. International Journal of Finance & Economics 19, 49–56.

Afonso, A., Furceri, D. and Gomes, P., 2012. Sovereign Credit Ratings and financial Markets Linkages: Application to European Data. Journal of International Money and Finance 31, 606–638.

Aizenman, J., 2013. The Eurozone Crisis: Muddling through on the Way to a more perfect Euro Union? Social Sciences 2, 221–233.

Ang, A. and Longstaff, F. A., 2013. Systemic Sovereign Credit Risk: Lessons from the US and Europe. Journal of Monetary Economics 60, 493–510.

Angelopoulou, E. and Balfoussia, H., Gibson, H. D., 2014. Building a financial Conditions Index for the Euro Area and selected Euro Area Countries: what does it tell us about the Crisis? Economic Modelling 38, 392–403.

Antonakakis, N. and Vergos, K., 2013. Sovereign Bond Yield Spillovers in the Euro Zone during the financial and Debt Crisis. Journal of International Financial Markets, Institutions and Money 26, 258–272.

Artis, M. J. and Zhang, W., 2001. Core and Periphery in EMU: A cluster Analysis. Economic Issues Journal Articles 6, 47–58.

Bafin, 2010. Ergebnisse der fuenften quantitativen Auswirkungsstudie zu Solvency II (QIS 5). Website, available online at: https://www.bafin.de/SharedDocs/Downloads/DE/Versicherer_Pensionsfonds/QIS/dl_qis5_ergebnisse_bericht_bafin.pdf?__blob=publicationFile&v=2; downloaded on 01/15/2017.

Basse, T., 2014. Searching for the EMU Core Member Countries. European Journal of Political Economy 34, S32–S39.

Basse, T. and Friedrich, M., 2008. Solvency II, Asset Liability Management, and the European Bond Market–Theory and empirical Evidence. Zeitschrift für die gesamte Versicherungswissenschaft 97, 155–171.

Basse, T., Friedrich, M. and Kleffner, A., 2012. Italian Government Debt and Sovereign Credit Risk: An empirical Exploration and some Thoughts about Consequences for European Insurers. Zeitschrift für die gesamte Versicherungswissenschaft 101, 571–579.

Basse, T., Friedrich, M., Kleffner, A. and von der Schulenburg, J.-M. G., 2014. Are Interest Rates too low? Empirical Evidence and Implications for German Life Insurers. Zeitschrift für die gesamte Versicherungswissenschaft 103, 31–43.

Bernoth, K. and Erdogan, B., 2012. Sovereign Bond Yield Spreads: A time-varying Coefficient Approach. Journal of International Money and Finance 31, 639–656.

Bernoth, K., von Hagen, J. and Schuknecht, L., 2012. Sovereign Risk Premiums in the European Government Bond Market. Journal of International Money and Finance 31, 975–995.

Bijlsma, M. and Vermeulen, R., 2016. Insurance Companies? Trading Behaviour during the European Sovereign Debt Crisis: Flight Home or Flight to Quality? Journal of Financial Stability 27, 137–154.

Blatt, D., Candelon, B. and Manner, H., 2015. Detecting Contagion in a multivariate Time Series System: An Application to Sovereign Bond Markets in Europe. Journal of Banking & Finance 59, 1–13.

Buechel, K., 2013. Do words matter? The Impact of Communication on the PIIGS'CDS and Bond Yield Spreads during Europe's Sovereign Debt Crisis. European Journal of Political Economy 32, 412–431.

Canarella, G., Miller, S. M. and Pollard, S. K., 2011. The global financial Crisis and stochastic Convergence in the Euro Area. International Advances in Economic Research 17, 315–333.

Chang, M. and Leblond, P., 2015. All in: Market Expectations of Eurozone Integrity in the Sovereign Debt Crisis. Review of International Political Economy 22, 626–655.

Claeys, P. and Vašíček, B., 2014. Measuring bilateral Spillover and testing Contagion on Sovereign Bond Markets in Europe. Journal of Banking & Finance 46, 151–165.

Cœuré, B., 2013. Outright monetary Transactions, one year on. American Economic Review 76, 72–81.

Costantini, M., Fragetta, M. and Melina, G., 2014. Determinants of Sovereign Bond Yield Spreads in the EMU: An optimal Currency Area Perspective. European Economic Re- view 70, 337–349.

De Backer, B., 2015. Decomposition of the dynamics of Sovereign Yield Spreads in the Euro Area. Economic Review pp. 54–75.

De Haan, L., Hessel, J. and van den End, J. W., 2014. Are European Sovereign Bonds fairly priced? The Role of modelling Uncertainty. Journal of International Money and Finance 47, 239–267.

De Haan, L. and Kakes, J., 2010. Are non-risk based Capital Requirements for Insurance Companies binding? Journal of Banking & Finance 34, 1618–1627.

Favero, C. A. and Missale, A., 2016. Contagion in the EMU – The Role of Eurobonds with OMTs. Review of Law & Economics 12, 555–584.

Ferreira, M. A. and Matos, P., 2008. The Colors of Investors' Money: The Role of institutional Investors around the World. Journal of Financial Economics 88, 499–533.

GDV, 2016. Statistical Yearbook of German Insurance 2016. Website, available on- line at http://www.en.gdv.de/wp-content/uploads/2016/11/Stat_Yearbook_2016.pdf; downloaded on 01/15/2017.

Giannetti, M. and Laeven, L., 2012. The Flight Home Effect: Evidence from the syndicated Loan Market during financial Crises. Journal of Financial Economics 104, 23–43.

Grinblatt, M. and Keloharju, M., 2000. The Investment Behavior and Performance of various Investor Types: A Study of Finland's unique Data Set. Journal of Financial Economics 55, 43–67.

Gruppe, M. and Lange, C., 2014. Spain and the European Sovereign Debt Crisis. European Journal of Political Economy 34, 3–8.

Harvie, C. and Pahlavani, M., 2006. Testing for Structural Breaks in the Korean Economy 1980-2005: An Application of the innovational Outlier and additive Outlier Models .

Iara, A. and Wolff, G. B., 2014. Rules and Risk in the Euro Area. European Journal of Political Economy 34, 222–236.

Kilponen, J., Laakkonen, H. and Vilmunen, J., 2012. Sovereign Risk, European Crisis Resolution Policies and Bond Yields .

Klose, J. and Weigert, B., 2014. Sovereign Yield Spreads during the Euro Crisis: Fundamental Factors versus Redenomination Risk. International Finance 17, 25–50.

Lane, P. R., 2012. The European Sovereign Debt Crisis. The Journal of Economic Perspectives 26, 49–67.

MacDonald, R., Sogiakas, V. and Tsopanakis, A., 2015. An investigation of systemic stress and interdependencies within the Eurozone and Euro Area countries. Economic Modelling 48, 52–69.

Maltritz, D., 2012. Determinants of Sovereign Yield Spreads in the Eurozone: A Bayesian Approach. Journal of International Money and Finance 31, 657–672.

Moro, B., 2014. Lessons from the European economic and financial great Crisis: A Survey. European Journal of Political Economy 34, 9–24.

Perron, P., 1990. Testing for a Unit Root in a Time Series with a changing Mean. Journal of Business & Economic Statistics 8, 153–162.

Perron, P., 1994. Trend, Unit Root and Structural Change in macroeconomic Time Series. Cointegration. Palgrave Macmillan, London, 1994. 113–146.

Perron, P. and Vogelsang, T. J., 1992. Testing for a Unit Root in a Time Series with a changing Mean: Corrections and Extensions. Journal of Business & Economic Statistics 10, 467–470.

Perron, P., 2006. Dealing with structural Breaks. Palgrave Handbook of Econometrics 1, 278–352.

Phillips, P. C. and Perron, P., 1988. Testing for a Unit Root in Time Series Regression. Biometrika 75, 335–346.

Randl, O. and Zechner, J., 2016. Sovereign Reputation and Yield Spreads: A Case Study on retroactive Legislation .

Schwarzbach, C., Kunze, F., Rudschuck, N. and Windels, T., 2014. Stock Investments for German Life Insurers in the current low Interest Environment: More Homework to do. Zeitschrift für die gesamte Versicherungswissenschaft 103, 45–63.

Sgherri, S. and Zoli, E., 2009. Euro Area Sovereign Risk during the Crisis .

Sibbertsen, P., Wegener, C. and Basse, T., 2014. Testing for a Break in the Persistence in Yield Spreads of EMU Government Bonds. Journal of Banking & Finance 41, 109–118.

Vogelsang, T. J. and Perron, P., 1998. Additional Tests for a Unit Root allowing for a Break in the Trend Function at an unknown Time. International Economic Review pp. 1073–1100.

Von Hagen, J., Schuknecht, L. and Wolswijk, G., 2011. Government Bond Risk Premiums in the EU revisited: The Impact of the financial Crisis. European Journal of Political Economy 27, 36–43.

Wegener, C., Basse, T., Kunze, F. and von Mettenheim, H.-J., 2016. Oil Prices and Sovereign Credit Risk of Oil producing Countries: An empirical Investigation. Quantitative Finance 16, 1961–1968.

Appendix

Table 3: Unit root break test GER - ESP; innovation outlier

Null Hypothesis: GERESP has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Innovational outlier

Break Date: 2010M03
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 1 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-2.726416	0.8203
Test critical values:	1% level	-4.949133	
	5% level	-4.443649	
	10% level	-4.193627	

Own calculations

Table 4: Unit root break test GER - ESP; additive outlier

Null Hypothesis: GERESP has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Additive outlier

Break Date: 2011M04
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 12 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-3.038404	0.6636
Test critical values:	1% level	-4.949133	
	5% level	-4.443649	
	10% level	-4.193627	

Own calculations

Table 5: Unit root break test GER - IT; innovation outlier

Null Hypothesis: GERIT has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Innovational outlier

Break Date: 2010M03
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 1 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-2.909122	0.7345
Test	1% level	-4.949133	
critical	5% level	-4.443649	
values:	10% level	-4.193627	

Own
calculations

Table 6: Unit root break test GER - IT; additive outlier

Null Hypothesis: GERIT has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Additive outlier

Break Date: 2010M02
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 1 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		--2.928872	0.7237
Test	1% level	-4.949133	
critical	5% level	-4.443649	
values:	10% level	-4.193627	

Own
calculations

Table 7: Unit root break test GER - AT; innovation outlier

Null Hypothesis: GERAT has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Innovational outlier

Break Date: 2008M08
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 3 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-3.868535	0.2063
Test	1% level	-4.949133	
critical	5% level	-4.443649	
values:	10% level	-4.193627	

Own calculations

Table 8: Unit root break test GER - AT; additive outlier

Null Hypothesis: GERAT has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Additive outlier

Break Date: 2007M11
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 0 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-2.923966	0.7266
Test	1% level	-4.949133	
critical	5% level	-4.443649	
values:	10% level	-4.193627	

Own calculations

Table 9: Unit root break test GER - NL; innovation outlier

Null Hypothesis: GERNL has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Innovational outlier

Break Date: 2008M02
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 2 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-4.539627	0.0388
Test	1% level	-4.949133	
critical	5% level	-4.443649	
values:	10% level	-4.193627	

Own calculations

Table 10: Unit root break test GER - NL; additive outlier

Null Hypothesis: GERNL has a unit root
Trend Specification: Intercept only
Break Specification: Intercept only
Break Type: Additive outlier

Break Date: 2008M10
Break Selection: Minimize Dickey-Fuller t-statistic
Lag Length: 2 (Automatic - based on Schwarz information criterion, maxlag=14)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-3.944137	0.1754
Test	1% level	-4.949133	
critical	5% level	-4.443649	
values:	10% level	-4.193627	

Own calculations

Chapter 3

Forecasting European interest rates in times of financial crisis — what insights do we get from international survey forecasts?

with Christoph Wegener, Kilian Bizer and Markus Spiwoks
contribution Frederik Kunze: 70%

Published:
Journal of International Financial Markets, Institutions and Money, Vol. 48, pp. 192–205, https://doi.org/10.1016/j.intfin.2017.01.005

Abstract

Interest rate forecasts are widely used in the international financial services industry. For decades, both practitioners and academic researchers question the quality and usefulness of forecasts. Survey predictions do not only deliver point forecasts but also allow to draw conclusions with regard to the variety of forecasts provided by professional analysts. We evaluate the quality of interest rate forecasts for the three months interbank rate in the UK (LIBOR) and Germany (EURIBOR) as well as the corresponding 10Y government bond yields using the root mean squared error as well as the Theil's U measure and also apply models of time series analysis (i.e. cointegration and causality analysis). Finally, we check for possible implications from uncertainty measures (i.e. High-Low-Spread of forecasts as well as forecast errors) and structural breaks. We are able to find some links to the real economy. Applying our methodological approach both to the UK and Germany we are able to draw conclusions with regard to the quality of international forecasts in times of uncertainty.

Keywords

Global financial crisis, survey forecasts, interest rate forecasts, forecast evaluation

JEL Classifications

G17, G01

Highlights

- Interest rate forecasts are important for financial market participants as well as monetary and political decision-makers.
- Survey predictions for interest rates are widely used but their accuracy may not be given as certain.
- The accuracy of interest rate forecasts both for the United Kingdom and Germany is investigated.
- High-Low-Spreads and Forecast Errors are used as dispersion measures to - examine the uncertainty of professional forecasters.
- In times of financial crisis financial market experts' uncertainty can be linked to real economic indicators.

J. Int. Financ. Markets Inst. Money 48 (2017) 192–205

Contents lists available at ScienceDirect

Journal of International Financial Markets, Institutions & Money

journal homepage: www.elsevier.com/locate/intfin

Forecasting European interest rates in times of financial crisis – What insights do we get from international survey forecasts?

Frederik Kunze [a,e], Christoph Wegener [b,c,*], Kilian Bizer [a], Markus Spiwoks [d]

[a] Faculty of Economic Sciences, Chair of Economic Policy and SME Research, University of Goettingen, Platz der Goettinger Sieben 3, 37073 Goettingen, Germany
[b] Ipag Business School, 184 Boulevard Saint-Germain, 75006 Paris, France
[c] Center for Risk and Insurance, Otto-Brenner-Straße 1, 30159 Hannover, Germany
[d] Faculty of Business Administration, Ostfalia University of Applied Sciences, Siegfried-Ehlers-Straße 1, 38440 Wolfsburg, Germany
[e] NORD/LB Norddeutsche Landesbank, Research/Economics, Friedrichswall 10, 30159 Hannover, Germany

ARTICLE INFO

Article history:
Received 19 May 2016
Accepted 19 January 2017
Available online 1 February 2017

JEL classification:
G17
G01

Keywords:
Global financial crisis
Survey forecasts
Interest rate forecasts
Forecast evaluation

ABSTRACT

Interest rate forecasts are widely used in the international financial services industry. For decades, both practitioners and academic researchers question the quality and usefulness of forecasts. Survey predictions do not only deliver point forecasts but also allow to draw conclusions with regard to the variety of forecasts provided by professional analysts. We evaluate the quality of interest rate forecasts for the three months interbank rate in the UK (LIBOR) and Germany (EURIBOR) as well as the corresponding 10Y government bond yields using the root mean squared error as well as the Theil's U measure and also apply models of time series analysis (i.e. cointegration and causality analysis). Finally, we check for possible implications from uncertainty measures (i.e. High-Low-Spread of forecasts as well as forecast errors) and structural breaks. We are able to find some links to the real economy. Applying our methodological approach both to the UK and Germany we are able to draw conclusions with regard to the quality of international forecasts in times of uncertainty.

1. Introduction

Interest rate forecasts are important for financial market participants' investment decisions. Since not all decision makers are willing or able to perform their own forecasts for the relevant term, forecast horizon or region, surveys collecting interest rate predictions are widely used. In addition to that, there exists empirical evidence for the general usefulness of survey forecasts (for a discussion of survey forecasts in general see for example Ang et al. (2007), Pesaran and Weale (2006), Schmeling and Schrimpf (2011) as well as Ince and Molodtsova (2017)). Predictions for interest rates are provided both for the short and the long end of the yield curve and for different countries. There already exists a vast literature dealing with the usefulness of the collected forecasts (i.e. both the individual predictions and the survey mean) for example with regard to rationality (see Friedman (1980), Simon (1989), Jongen and Verschoor (2008) as well as Chortareas et al. (2012) and more recently Miah et al. (2016)), unbiasedness (Hafer and Hein (1989) as well as Mitchell and Pearce (2007) and Miah et al. (2016)), efficiency (see for example Hafer and Hein (1989)) and accuracy (see for example Kolb and Stekler (1996) as well as Greer (2003)).

* Corresponding author at: Ipag Business School, 184 Boulevard Saint-Germain, 75006 Paris, France.
E-mail addresses: cw@kvw-hannover.de, c.wegener@ipag.fr (C. Wegener), bizer@wiwi.uni-goettingen.de (K. Bizer), m.spiwoks@ostfalia.de (M. Spiwoks).

http://dx.doi.org/10.1016/j.intfin.2017.01.005
1042-4431/

Furthermore, summing up the results from the relevant literature would inevitably lead to the disillusioning conclusion that especially in times of financial crisis forecasts from professional analysts seem to perform fairly badly. This puts the usefulness of into question. However, in the context of crisis events dispersion measures have been investigated empirically. Dispersion measures may act as an indicator for uncertainty. Hence, a higher level of uncertainty might serve as a bellwether for dawning crisis events.

The contribution of this paper to the existing literature is threefold. Firstly, a long term assessment of survey forecasts for the three months interbank rate for two major European economies (i.e. the United Kingdom and Germany) and for the 10 year government bond yield will be performed using standard evaluation methods. Secondly, we will use more sophisticated methods of time series analysis to investigate the quality of survey forecasts (i.e. the survey mean) in an international context. Finally, and most important in the context of the global financial crisis, we try to get insights from uncertainty measures (i.e. the dispersion of the collected individual forecasts as well as the forecast errors). We try to find crisis related structural changes in the time series of dispersion measures and forecast errors. Since we understand possible structural breaks to be crisis-related the research question behind this approach points in the direction of early warning indicators or at least some kind of bellwether or discretional warning sign. In order to find evidence we link the uncertainty measures to real economic indicators using Granger causality analysis. By applying this methodological approach both to the UK and to Germany we try find parallels and want to rule out possible country specific implications.

To the best of our knowledge this is the first study to address the relevant question of combining uncertainty of interest rate forecast with real economic data in the context of the global financial crisis focusing on Germany and the United Kingdom. In this sense, we both deliver useful results for the forecast evaluation literature and fill a research gap in the context of possible crisis related relationships between forecaster uncertainty and real economic activity. One major advantage of our approach lies in the combination of easy to interpret forecast accuracy measures with time series analysis for a two country data set. In addition to that, we are able to cross link the financial market variables to the real economy. For political decision makers uncertainty measures can under certain circumstances be used as crisis indicators.

The structure of this paper is as follows. Chapter 2 provides a short literature overview whereas in chapter 3 the relevant data set of forecasts will be introduced. In chapter 4 we will briefly present the applied forecast evaluation methods whereas the empirical results will be discussed in chapter 5. Finally, we will shortly present the idea of uncertainty measures in chapter 6. In the same section we will also present some empirical evidence for crisis related structural breaks for the discussed uncertainty measures and check for implications from the uncertainty measures for real economic activity. In chapter 7 we will discuss possible economic implications. Chapter 8 concludes the paper.

2. Literature review

Not least due to the relevance of the accuracy of forecasts for decision makers the literature dealing with the evaluation of economic as well as financial market forecasts from professional analysts dates back for several decades. This is also the case for the research dealing with the evaluation of interest rate forecasts (see for example Friedman (1980), Belongia (1987), Spiwoks et al. (2008), as well as Hafer and Hein (1989)). One important field of empirical research has been the evaluation of survey forecasts. Dependent on the interest rates being predicted by professional analysts there do exist a lot of well-known and often cited surveys (dealing mainly with the US economy).

Greer (2003) for example tested the directional accuracy of long-term interest rate forecasts issued by The Wall Street Journal. Baghestani (2006) did assess the accuracy of survey forecasts for 10Y US Treasury bond yields delivered to the Survey of Professional Forcasters (SPF) collected by the Federal Reserve Bank of Philadelphia. Although this survey provides one of the longest histories of forecasts its applicability as data source to evaluate crisis related changes has at least to be questioned because the survey is performed on a quarterly basis. Spiwoks et al. (2008) did also investigate US interest rate forecasts using survey data collected by Consensus Economics Inc. One advantage of these surveys lies in the fact that apart from the US, there exist professional analysts' predictions for interest rates for various countries.

Kunze et al. (2014) did examine interest forecasts for the three months EURIBOR and also checked for crisis related changes in the long term relationship of the forecasts and the actual interest rate. Chortareas et al. (2012) evaluated interest rate survey forecasts from Consensus Economics Inc. for the UK three months interbank rate as well as 10Y government bond yield with regard to the rational expectations hypothesis also dealing with the aspects of both monetary policy actions and the central banks communication policy (see Stillwagon (2015) for a recent study testing the expectations hypothesis with survey forecasts). Despite the different focus of the paper the work of the authors is of high relevance in the context of our paper and the underlying data. Not only market participants but also political decision makers have to decide under uncertainty. The results of the mentioned studies are of high relevance for our analysis since we want to put the findings of the authors in an international context.

In times of financial crisis the concept of measuring uncertainty (which we assume to rise during crisis events) becomes more important when using professional analysts' predictions for interest rates during the process of decision making. It has to be taken into account that long run relationships may not hold in times of financial crisis. For example, Sibbertsen et al. (2014) tested for structural changes in the spreads of interest rates and Wegener et al. (2016) used these results in order to forecast German and French government bond yields. In periods of high uncertainty forecasters might alter their predictions not to the same extent. Hence, by solely focusing on the mean of the survey forecast decision makers might be aware of a high possibility of pronounced movements in future interest rates (at least they might be aware that there exists consensus in which direction interest rates might move). But the information whether this mean predictions stem from a broad variety of individual point forecasts would be ignored. However, the degree of uncertainty might be needed as much more crucial information than the mean expectations regarding the future interest rates.

Measuring uncertainty is important for decision makers for at least two reasons. Firstly, the simple awareness to be in a period of high financial market uncertainty might work as a warning signal for decision makers. Secondly, rising uncertainty with regard to the future movements of interest rates might also work as some kind of early warning indicator for possible economic downturns and hence the direction of movement for real economic indicators. Dealing with the concepts of uncertainty it has to be stated that there do exist a lot of different methods to measure uncertainty. A widely used measure for uncertainty is the dispersion of individual forecasts. Following Givoly and Lakonishok (1984) the forecast dispersion for earnings is at least perceived by financial market participants to be a good proxy for the uncertainty of a firms' actual future earnings. For example Bloom (2009) did use the volatility of a stock market index as a measure of uncertainty. The advantage of this approach lies in the high data frequency, which almost allows a real time monitoring. Gupta et al. (2014) did investigate US economic policy uncertainty and financial stress with regard to their ability to predict US equity premium. In the context of survey forecasts for example Rich et al. (1992) discussed the relationship between measures of forecast dispersion and forecast uncertainty from survey data whereas Bomberger (1996) did emphasize the need to differentiate between individual uncertainty and cross sectional uncertainty. More recently the relationship between forecast uncertainty and forecast disagreement has been inter alia discussed by Lahiri and Sheng (2010).

With regard to the different measures of uncertainty (which we do not only understand as one of the major sources of forecast errors but also as indication for possible crisis events) one of the most relevant papers stems from Zarnowitz (1992). Although the author came to the conclusion that the usefulness of dispersion measures (i.e. the difference between the highest and the lowest provided forecast at one point in time) as an indicator of uncertainty has to put into question, their remarks are of high relevance for our analysis not only because Zarnowitz (1992) further discusses whether the interpersonal dispersion of forecasts (measured by the standard deviation of individual forecasts) is an acceptable proxy for the dispersion of intrapersonal predictive probabilities uncertainty. As a matter of fact, we are not interested in the question whether interpersonal uncertainty is a good proxy for intrapersonal uncertainty, but we want to check for possible implications from interpersonal uncertainty, which we can easily derive from survey forecasts.

3. Data

In this paper we investigate interest rate forecasts for two major European economies: The United Kingdom (UK) and Germany (GER). For both countries the three months interbank rate (from now on 3M) as well as the yield of 10 Year Government bonds (from now on 10Y) provided by Consensus Economics Inc. will be examined. The surveys for UK and GER are conducted with two forecast horizons: A three months as well as a twelve months ahead forecast. In fact, the actual forecast horizons are four respectively thirteen months, because the forecasters have to provide their predictions at the beginning of the month (see for example Spiwoks et al., 2008).

The UK sample under investigation does range from December 1993 to December 2014. The actual time series are the 3 month UK interbank rate (3M) and the 10 year UK government bond yields (10Y). Survey forecasts for both time series with a 4M respectively 13M forecast horizon are denoted as 3M4M UK and 3M13M UK respectively 10Y4M UK and 10Y13M UK. For each forecast time series the highest, the lowest as well as the naive forecast are also available.

The GER sample under investigation ranges from January 2000 to August 2014. The shorter time horizon has been chosen to rule out any structural breaks due to the introduction of the Euro. Apart from the time horizon the data set for Germany follows a similar structure to the UK data set.

A summary of the abbreviations both for UK and for GER are given in Table 1 below.

4. Methodologies of forecast evaluation

We start the empirical analysis of the UK respectively GER interest rate forecasts with simple evaluation measures. Although the root mean squared error (RMSE) is a crude method in the context of forecast evaluation it is widely used by

Table 1
Abbreviations.

Time series	Abbreviation
UK 4 months forecast for the 3 months interbank rate	3M4M UK
UK 13 months forecast for the 3 months interbank rate	3M13M UK
UK 4 months forecast for the 10Y government bond yield	10Y4M UK
UK 13 months forecast for the 10Y government bond yield	10Y13M UK
GER 4 months forecast for the 3 months interbank rate	3M4M GER
GER 13 months forecast for the 3 months interbank rate	3M13M GER
GER 4 months forecast for the 10Y government bond yield	10Y4M GER
GER 13 months forecast for the 10Y government bond yield	10Y13M GER

practitioners in the financial service industry due to the straightforward interpretation and the relevance in statistical methodologies (see for example Hyndman and Koehler (2006)). Since this measure is widely known and in order to preserve space the formula will not be presented in detail. This also holds for the Theil's U which we use to further analyze the forecast performance (see Theil, 1992) by comparing the forecast with a simple naive prediction.

We apply both forecast accuracy measures to get a first idea of the quality of the 4 months respectively 13 months ahead predictions for the three months interbank rate as well the 10Y government bond yields for UK and GER.

Furthermore, we use the widely applied Diebold Mariano (DM) Test (see Diebold and Mariano, 1995) to check whether the mean forecasts' accuracy is different from the accuracy of the simple naive forecast. One advantage of the DM Test lies in the fact that we are able to draw conclusions with regard to statistical significance.

We start the empirical analysis by checking for the order of integration both for the eight forecast time series as well as for the four actual time series. Instead of using the commonly used ADF-test (see Dickey and Fuller, 1979) we test for possible unit roots applying the procedure from Phillips and Perron (1988), since this procedure tends to be more robust. For example Cheung and Chinn (1998) did point out that for FX forecasts to be rational the predictions should be of the same order of integration like the predicted time series. Kunze et al. (2014) applied this idea to survey forecasts for the three months EURIBOR whereas Schwarzbach et al. (2012) have evaluated this most basic requirement for the usefulness of survey forecasts for the 10Y German government bond yields with a forecast horizon of thirteen months and with a different time frame.

Also following Cheung and Chinn (1998) as well as Schwarzbach et al. (2012) and Kunze et al. (2014) after deriving the order of integration it will be tested for long run relationships between the eight interest rate time series and the corresponding survey means to check whether there do exist long run relationships between the actual interest rates and the mean forecasts which is also a basic requirement for rational forecasts. To check for the existence of long run relationships we employ the often used Johansen procedure (see Johansen, 1991) applying the trace test statistic.

Two time series with the same order of integration are said to be cointegrated if they share a common stochastic trend. In this case there must be at least unidirectional Granger causality. If the forecasters do their job in a proper way both series – the prediction and the actual time series – should have the same stochastic properties and thus they should be cointegrated. Because of this we apply methodologies from the field of cointegration.

We conclude by checking for causal relationships between the actual time series and the corresponding forecasts and again follow the approach applied by Schwarzbach et al. (2012) for the 10Y German government bond yield and also check for causal relationships (i.e. Granger causality) by using impulse response functions with orthogonal (see Sims, 1980) impulses and a length of the impulse responses of 30 month ahead.

5. Evaluation results

Table 2 shows the Theil's U results for the survey means for the UK data as well as the GER data.

It seems that the survey means are slightly better compared to the naive forecasts as reported in Table 2. However, this result is only statistically significant for the long run interest rates as reported in Table 3 on a level of 10%. As Table 4 below shows the time series for the actual interest rates as well as the survey means both for the UK data and for the GER data are non-stationary on levels but stationary in first differences (d(-)). Hence, both the actual interest rates and the mean forecast are I(1) and share the same order of integration. Accordingly, all four mean forecasts for each country fulfill the most basic requirement for rational forecasts.

Checking for long run relationships (i.e. cointegration) between the actual interest rate as well as the mean forecast for the eight pairs of time series we find evidence for a cointegration relationship for each pair of forecast and actual time series. Hence, it has been statistically validated that also the second prerequisite has been fulfilled (see Tables 5 and 6 below). The null hypothesis of no cointegration has to be rejected for all eight cases.

Although we found evidence that the time series share the same order of integration and statistically verified that a long run relationship exists both for UK and for GER, this does not necessarily mean that the survey mean forecasts are accurate respectively useful. To be useful the forecasts should deliver some relevant information for the future path of the predicted interest rates. To check for Granger causality between the eight corresponding pairs of time series we estimate eight vector error correction models (VECM) and check for Granger causality using impulse response functions (IRF). Some preliminary considerations should be made: As for example Basse and Reddemann (2010) discussed, the ordering of the variables in the

Table 2
Theil's U.

	UK			GER		
	RMSE Mean	RMSE Naive	Theil's U	RMSE Mean	RMSE Naive	Theil's U
3M4M	0.6325	0.5048	1.2529	0.3466	0.4469	0.7755
3M13M	1.0207	0.8869	1.1507	1.0681	1.2670	0.8430
10Y4M	0.6256	0.6337	0.9872	0.2173	0.3763	0.5775
10Y13M	1.2386	1.4177	0.8737	0.4761	0.6712	0.7093

Table 3
P-values Diebold Mariano test.

DM test	Survey mean UK	Survey mean GER
3M4M	0.8708	0.2374
3M13M	0.3702	0.3111
10Y4M	0.0005	0.0081
10Y13M	0.0940	0.0064

Table 4
P-values of the PP tests for UK and GER.

	Actual	*d* (actual)	Mean	*d* (mean)
10Y4M UK	0.06	0.01	0.25	0.01
10Y13M UK	0.06	0.01	0.32	0.01
3M4M UK	0.40	0.01	0.38	0.01
313M UK	0.40	0.01	0.42	0.01
10Y4M GER	0.10	0.01	0.08	0.01
10Y13M GER	0.10	0.01	0.15	0.01
3M4M GER	0.71	0.01	0.65	0.01
313M GER	0.71	0.01	0.60	0.01

Table 5
Results of the Johansen procedure for the UK.

	Test statistic	10%	5%	1%
10Y4M UK				
$r \leqslant 1$	0.33	6.50	8.18	11.65
$r = 0$	82.66	12.91	14.90	19.19
10Y13M UK				
$r \leqslant 1$	0.24	6.50	8.18	11.65
$r = 0$	19.24	12.91	14.90	19.19
3M4M UK				
$r \leqslant 1$	0.73	6.50	8.18	11.65
$r = 0$	97.00	12.91	14.90	19.19
3M13M UK				
$r \leqslant 1$	0.64	6.50	8.18	11.65
$r = 0$	29.81	12.91	14.90	19.19

Table 6
Results of the Johansen procedure for Germany.

	Test statistic	10%	5%	1%
10Y4M GER				
$r \leqslant 1$	1.09	6.50	8.18	11.65
$r = 0$	27.82	12.91	14.90	19.19
10Y13M GER				
$r \leqslant 1$	0.91	6.50	8.18	11.65
$r = 0$	27.57	12.91	14.90	19.19
3M4M GER				
$r \leqslant 1$	2.91	6.50	8.18	11.65
$r = 0$	98.41	12.91	14.90	19.19
3M13M GER				
$r \leqslant 1$	2.62	6.50	8.18	11.65
$r = 0$	27.34	12.91	14.90	19.19

model plays an important role in applied econometrics. Since we are interested in the usefulness of the interest forecasts we put the time series containing the mean forecasts first. The confidence intervals (CI) with a 95% level have been calculated via bootstrapping (see Efron and Tibshirani, 1994). The results are illustrated from Figs. 1–16 below.

A significant response from the act time series to the impulse (i.e. a positive shock) from the corresponding mean time series (or vice versa) would be indicated by the response function by a move of both the impulse function as well as the two confidence bands above or below the zero line. The results from the impulse response analysis imply that only for the four months forecast of the UK 10Y government bond yield (i.e. mean1 respectively act1) empirical evidence for bi-directional Granger causality can be found. The lower segment of Fig. 1 shows that the response function (solid line) indicates a significant negative response, because in addition to the response function both bounds of the confidence band (dotted line) fall

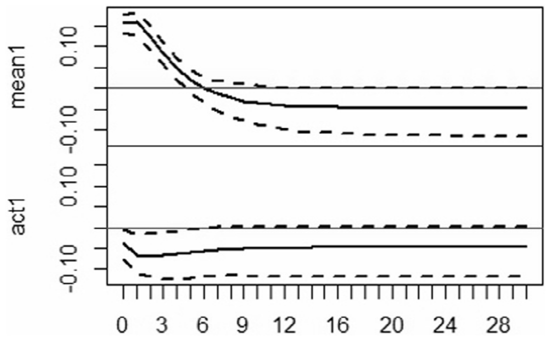


Fig. 1. Impulse response function from mean1 (10Y4M UK) to mean1 (upper curve) and to act1 (curve below).

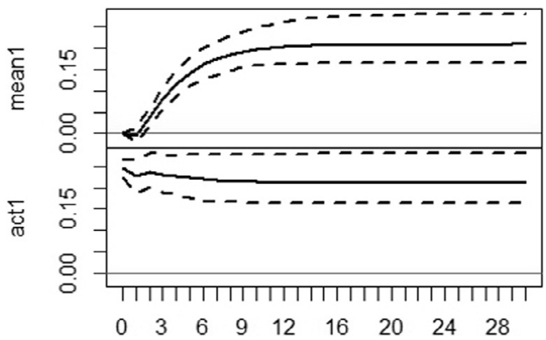


Fig. 2. Impulse response function from act1 (10Y4M UK) to mean1 (upper curve) and to act1 (curve below).

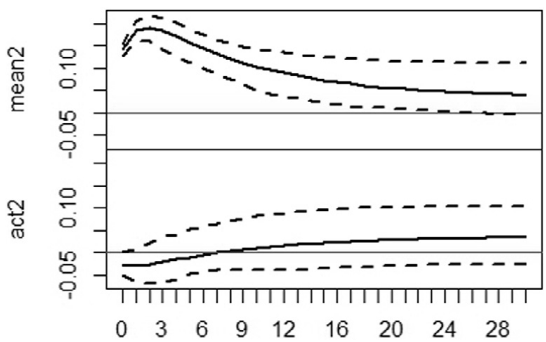


Fig. 3. Impulse response function from mean2 (10Y13M UK) to mean2 (upper curve) and to act2 (curve below).

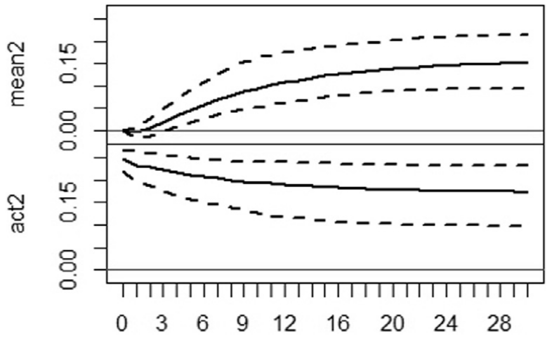


Fig. 4. Impulse response function from act2 (10Y13M UK) to mean2 (upper curve) and to act2 (curve below).

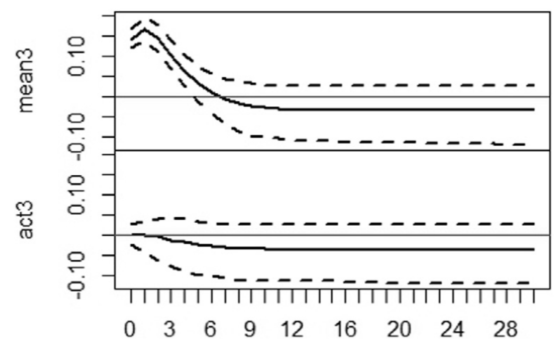


Fig. 5. Impulse response function from mean3 (3M4M UK) to mean3 (upper curve) and to act3 (curve below).

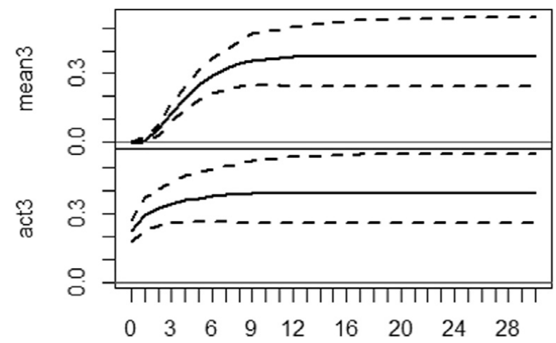


Fig. 6. Impulse response function from act3 (3M4M UK) to mean3 (upper curve) and to act3 (curve below).

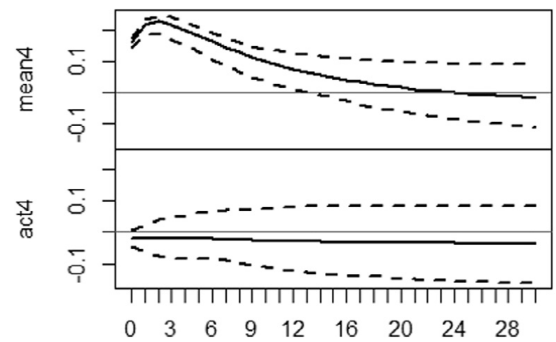


Fig. 7. Impulse response function from mean4 (3M13M UK) to mean4 (upper curve) and to act4 (curve below).

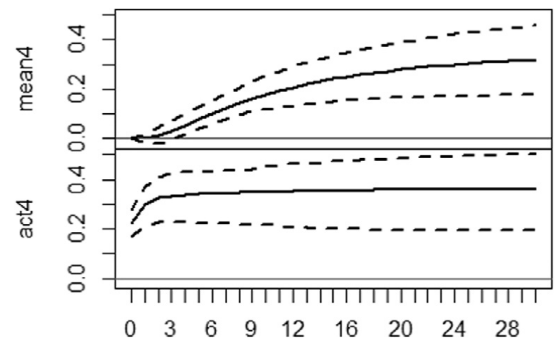


Fig. 8. Impulse response function from act4 (3M13M UK) to mean4 (upper curve) and to act4 (curve below).

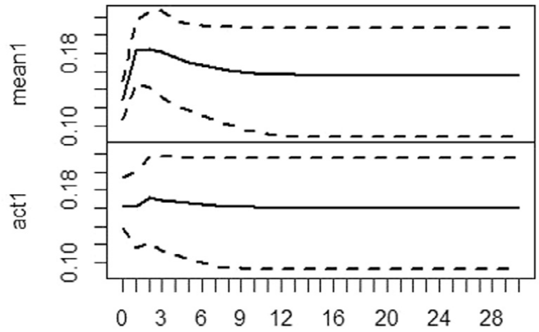


Fig. 9. Impulse response function from mean1 (10Y4M GER) to mean1 (upper curve) and to act1 (curve below).

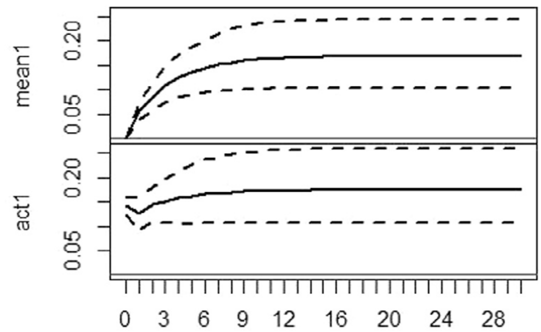


Fig. 10. Impulse response function from act1 (10Y4M GER) to mean1 (upper curve) and to act1 (curve below).

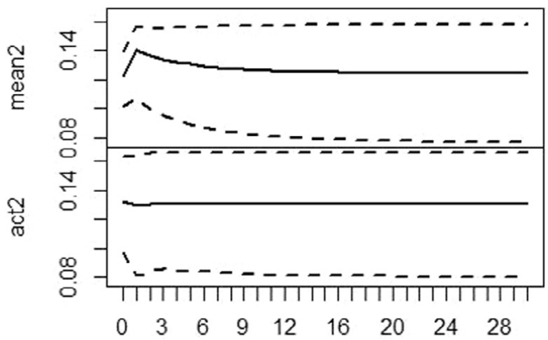


Fig. 11. Impulse response function from mean2 (10Y13M GER) to mean2 (upper curve) and to act2 (curve below).

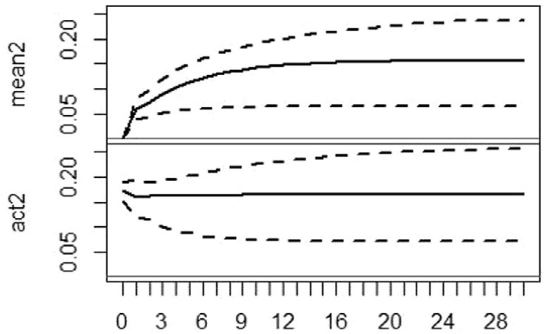


Fig. 12. Impulse response function from act2 (10Y13M GER) to mean2 (upper curve) and to act2 (curve below).

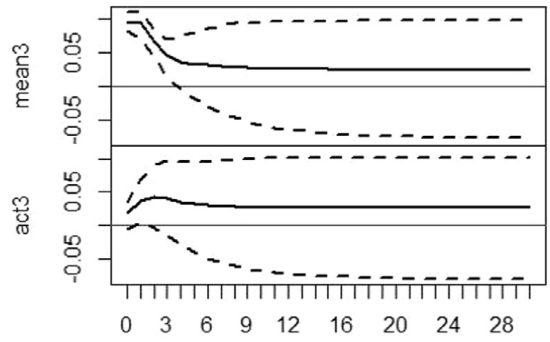


Fig. 13. Impulse response function from mean3 (3M4M GER) to mean3 (upper curve) and to act3 (curve below).

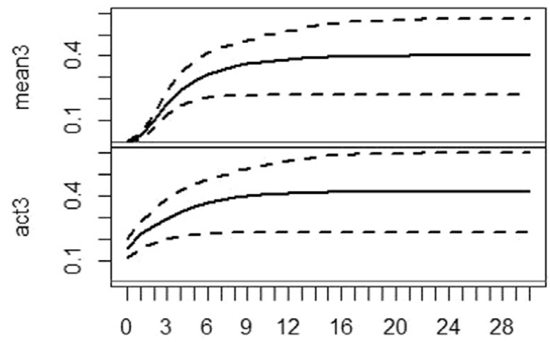


Fig. 14. Impulse response function from act3 (3M4M GER) to mean3 (upper curve) and to act3 (curve below).

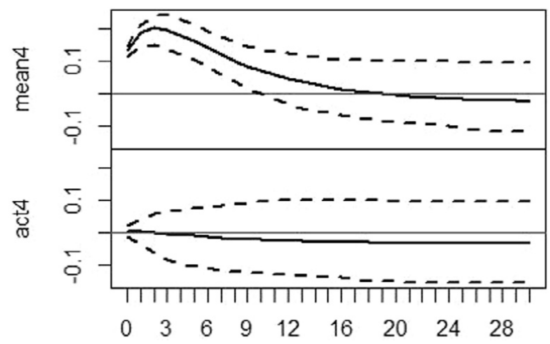


Fig. 15. Impulse response function from mean4 (3M13M GER) to mean4 (upper curve) and to act4 (curve below).

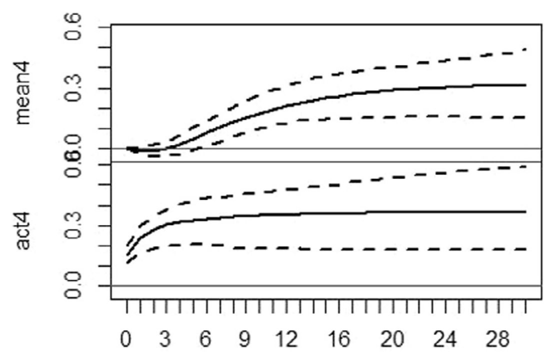


Fig. 16. Impulse response function from act4 (3M13M GER) to mean4 (upper curve) and to act4 (curve below).

below the zero line. However, the interpretation of this causal relationship is not unambiguous, because the IRF results indicate that a positive shock from the survey mean causes an instantaneous negative response.

Interestingly (as can easily be seen from the upper segments in the Figs. 3–16), the Granger causality running from the actual interest rates and the corresponding survey mean time series results in a positive response. This clearly is more corresponding to the economic interpretation in the context of this paper. Summing up the results for the two different countries leads to the conclusion that from a time series analysis perspective all survey mean forecasts despite the UK 10Y4M forecast are not delivering any relevant information regarding the future movement of the 10Y government bond yield and the three months interbank rate (i.e. there is only evidence for unidirectional Granger causality running from the actual interest rate to the survey mean time series). In addition to that, the only exception (i.e. the four months forecast for the 10Y GILT yield) is not corresponding to economic theory.

6. Examining possible implications from dispersion measures and forecast errors

The evaluation results presented in Chapter 5 above put the forecasting performance of survey forecasts into question. With regard to the survey mean (both for the case of the UK data and the GER data) we did not find strong evidence for the usefulness of professional forecasts. Anyhow, bearing in mind the exceptional market reactions due to past financial respectively economic crisis we want to check for possible insights from forecast errors and dispersion measures in the context of international crisis events. Both the dispersion of forecasts (i.e. the difference between the highest and the lowest forecast, from now on High-Low-Spread, HLS) and the forecast errors (i.e. the difference between the mean forecast and the actual interest rate) seem to be remarkably unstable in times of financial turmoil.

This should not be a surprise at all, since both measures can be seen as indicators of uncertainty (see for example Kunze et al., 2014 who investigated the HLS of professional forecast for the three months EURIBOR in the context of the financial crisis). Using the forecast errors we are able to observe the uncertainty as deviation from the actual interest rate. However, it is only an ex post measure of uncertainty which is subject to a substantial time lag (i.e. four respectively 13 months).

Because of that, we analyze both the forecast errors and the HLS for the interest rates under investigation. The abbreviations of the relevant time series are presented in Table 7 below. If the HLS proved to be a good measure of uncertainty it would be superior to forecast errors at least with regard to the timing, because we do not have to compare the HLS for a specific forecast with the actual interest rate in four respectively 13 months. To check for possible implication from the two described uncertainty measures we will check for structural breaks in the 16 relevant time series.

We want to get to know crisis related implications from the forecast errors as well as the HLS. If uncertainty is to rise in times of financial crisis these time series might serve as some kind of warning sign for political decision makers or market participants.

As a matter of fact, the forerunning characteristic of interest rates for the real economy has been widely discussed (see for example Bernanke (1990), Bernanke and Blinder (1992), Bernard and Gerlach (1998), Estrella et al. (2003) and more recently Saar and Yagil (2015) who examined the usefulness of government and corporate yield curves to predict economic growth and stock market reactions in the United Kingdom, Europe and Japan). Stock and Watson (1989) also argued that interest rates deliver useful information for real economic variables. Using again the concept of Granger causality (see Granger (1969)), we want to check for Granger causal relationships between economic indicators and the outlined uncertainty measures (i.e. FECONS and HLS respectively). Since GDP growth is only available on a quarterly basis, this indicator is not applicable in the context of this paper. Being available on a monthly basis and covering a large part of real economic activity in the United Kingdom as well as in Germany we chose to check for causal relationships between the uncertainty measures and the index of Industrial production (ip). In addition to that we investigate the relationship between the uncertainty measures and

Table 7
Abbreviations for Forecast errors and High-Low-Spreads.

Description	Abbreviation
UK Forecast errors 4 months forecast for the 3 months interbank rate	FECONS1 UK
UK Forecast errors 13 months forecast for the 3 months interbank rate	FECONS2 UK
UK Forecast errors 4 months forecast for the 10Y government bond yield	FECONS3 UK
UK Forecast errors 13 months forecast for the 10Y government bond yield	FECONS4 UK
UK High-Low-Spread 4 months forecast for the 3 months interbank rate	HLS1 UK
UK High-Low-Spread 13 months forecast for the 3 months interbank rate	HLS2 UK
UK High-Low-Spread 4 months forecast for the 10Y government bond yield	HLS3 UK
UK High-Low-Spread 13 months forecast for the 10Y government bond yield	HLS4 UK
GER Forecast errors 4 months forecast for the 3 months interbank rate	FECONS1 GER
GER Forecast errors 13 months forecast for the 3 months interbank rate	FECONS2 GER
GER Forecast errors 4 months forecast for the 10Y government bond yield	FECONS3 GER
GER Forecast errors 13 months forecast for the 10Y government bond yield	FECONS4 GER
GER High-Low-Spread 4 months forecast for the 3 months interbank rate	HLS1 GER
GER High-Low-Spread 13 months forecast for the 3 months interbank rate	HLS2 GER
GER High-Low-Spread 4 months forecast for the 10Y government bond yield	HLS3 GER
GER High-Low-Spread 13 months forecast for the 10Y government bond yield	HLS4 GER

Table 8
Results PP-test economic indicators (levels and first differences).

	FECONS1	FECONS2	FECONS3	FECONS4
PP-Test (UK)	0.01	0.01	0.01	0.07
PP-Test (GER)	0.01	0.37	0.01	0.01
	HLS1	HLS2	HLS3	HLS4
PP-Test (UK)	0.01	0.01	0.01	0.01
PP-Test (GER)	0.01	0.01	0.01	0.01
	ip	dip	ue	due
PP-Test (UK)	0.27	0.01	0.36	0.01
PP-Test (GER)	0.50	0.01	0.92	0.01

Table 9
Summary breakpoints.

	BIC	Timing of break				
FECONS1 UK	0					
FECONS2 UK	3	1996(12)	2000(1)	2008(7)		
FECONS3 UK	3	1998(9)	2003(7)	2008(10)		
FECONS4 UK	4	2000(12)	2004(1)	2008(10)	2011(11)	
HLS1 UK	4	1996(12)	2000(9)	2008(10)	2011(11)	
HLS2 UK	3	1996(12)	2000(7)	2009(4)		
HLS3 UK	3	1996(12)	2007(9)	2010(10)		
HLS4 UK	4	1996(12)	2004(8)	2008(10)	2011(11)	
FECONS1 GER	3	2005(11)	2008(10)	2010(12)		
FECONS2 GER	5	2002(2)	2004(4)	2006(6)	2008(11)	2011(1)
FECONS3 GER	3	2002(5)	2006(12)	2009(8)		
FECONS4 GER	4	2002(7)	2005(10)	2009(8)	2012(6)	
HLS1 GER	3	2004(1)	2007(12)	2010(2)		
HLS2 GER	4	2002(6)	2006(9)	2009(3)	2011(5)	
HLS3 GER	3	2002(10)	2006(6)	2009(5)		
HLS4 GER	2	2008(5)	2010(7)			

the unemployment rate (ue). One major advantage of these economic indicators is the straightforward interpretation which also allows to compare the cause of the economic crisis in both countries.

Once again, we start our empirical analysis by deriving the order of integration for the time series under investigation to rule out the possibility of spurious results. As a matter of fact, if the time series of the uncertainty measures as well as the economic indicators would be of the same order of integration we would have to check for cointegrating relationships. In Table 8 below the results of the corresponding PP-Tests are presented.

For almost all 16 uncertainty measures the PP-Test delivers empirical evidence (1% significance level) that the uncertainty measures are integrated of order zero. The thirteen months forecast for the three 10Y UK government bond yield (i.e. FECONS4 UK) appears only to be I(0) on a 10% significance level. More surprisingly in the case of the Germany the thirteen months forecast for the 3M interbank rate (FECONS2 GER) seems to be stationary. Following the empirical results in Table 8 above the economic indicators ip and ue are I(1) in levels respectively I(0) in first differences. Summarizing the results we conclude that a further investigation of long run relationships is not indicated in general. For the case of the FECONS2 GER time series we did apply the Johansen Test to rule out a cointegrating relationship with the economic variables. To rule out any spurious results we will not investigate this time series when checking for Granger causal relationships. As a matter of fact following the Johansen Test for cointegration we found empirical evidence for a long run relationship between FECONS2 GER and ue, but zero cointegration relationships between FECONS2 GER and ip.

We go on by testing for structural breaks in the time series of forecast errors as well as the HLS following the approach outlined by Zeileis et al. (2003). The authors present an applied framework to test for structural breaks with unknown timing. In addition to that the procedure outlined by Zeileis et al. (2003) also allows to check for the occurrence of several breaks in the time series under investigation. The time series described above cover a time frame from December 1993 to December 2014 (and hence more than 20 years) for the UK and from January 2000 to August 2014 in the case of Germany. This is important in the context of this paper. Since we understand the forecast errors as well as the HLS as uncertainty measures, shifts in these 16 time series might point to crisis related increases in uncertainty. Hence, the timings of the structural breaks are of interest. Following Zeileis et al. (2003) as well as Zeileis (2006) we want to check whether the mean of the time series (i.e. FECONS and HLS) changes over time using an OLS-Based CUSUM test (see Ploberger et al., 1989; Zeileis et al., 2003) as well as a sequence of F-statistics (see Zeileis et al., 2003). After checking for the existence of structural breaks we again follow Zeileis et al. (2003) and apply a dynamic programming algorithm from Bai and Perron (2003) to derive the corresponding break dates.

Table 10
P-values and number of lags Granger causality test. dip refers to the first differences of industrial production and due refers to first differences to the unemployment rate.

	FECONS1	Lags	FECONS2	Lags	FECONS3	Lags	FECONS4	Lags
UK dip			0.43	1	0.05	2	0.08	2
GER dip	0.00	3			0.20	1	0.25	1
UK due			0.91	1	0.18	2	0.07	2
GER due	0.10	2			0.65	1	0.07	1
	HLS1	Lags	HLS2	Lags	HLS3	Lags	HLS4	Lags
UK dip	0.83	1	0.01	1	0.96	1	0.07	1
GER dip	0.68	1	0.09	1	0.02	1	0.36	1
UK due	0.07	2	0.36	2	0.14	1	0.47	2
GER due	0.66	1	0.32	2	0.22	1	0.91	1

Following the results of the breakpoint analysis the only time series where no statistical evidence for a structural break has been found is the forecast error of the three months forecast for the UK 10Y government bond yield (i.e. FECONS1 UK). Because of that, we also exclude this time series from the Granger causality tests below. In order to preserve space the results of the breakpoint analysis are available on request from the authors. A summary of the results from the breakpoint analysis is presented in the Table 9 below. The most important result from the breakpoint analysis in the context of this paper is the timing of the structural breaks.

Since the timing of a substantial part of the breakpoints indicates mean shifts in the uncertainty measures around (financial) crisis events respectively during periods of crisis (e.g. October 2008) we go on by trying to find causal relationships to the real economic indicators and once again apply the concept of Granger causality. Due to the fact that we did not find any long run relationships for the remaining 14 time series we use a simple vector auto regressive model (VAR) to test for Granger causality between the forecast errors respectively HLS and the first differences of the industrial production time series (dip) and the first differences of the unemployment rate (due).

The first differences of ip and ue have been chosen because non stationary variables might lead to distorted results. We estimated 14 VAR-models and checked for Granger causality whereas the FECONS- respectively HLS-time series were the cause variables (see Pfaff, 2008). For each model the lag length has been derived using the SC criterion proposed by Schwarz (1978). The causality analysis tested against the null hypothesis of no Granger causality. The results (lag length and corresponding p-values) are presented in Table 10 below.

The estimations above deliver mixed results. With regard to the significance level we found some empirical evidence for Granger causal relationships. On the 1% level FECONS2 GER Granger causes changes in the German ip (dip) and HLS2 UK Granger causes changes in the UK industrial production (dip). On the 5% significance level FECONS3 UK Granger causes dip whereas HLS2 GER Granger causes dip. These results are encouraging in the sense that changes in the uncertainty measures may act as some kind of warning sign in financial respectively economic crisis at least with regard to the industrial production. The evidence with regard to the unemployment rate is less encouraging, however.

7. Policy implications

Dealing with the causes and consequences of financial crisis events is a demanding task for financial market participants as well as political decision makers. Additionally, the global financial and economic crisis did visualize quite impressively that crisis events and the consequences thereof are not regionally restricted and hence have to be analyzed in an international context. For example, monetary policy decisions of the European Central Bank (ECB) have to be considered by the Bank of England's (BoE) monetary policy committee. Not surprisingly and as an answer to the economic downturn in the aftermath of the global financial crisis both the BoE's and ECB's decision makers did cut benchmark interest rates several times to stimulate economic activity. In addition to that, it has to be considered that the long end of the yield curve for various countries has also been affected by the rising risk aversion of financial market participants. In this context Germany and the United Kingdom might be considered as safe havens. During the financial crisis both the 10Y German government bond yield the 10Y GILT yields did fall substantially due to flight-to-quality effects. On the other hand, government yields for countries like Portugal, Italy and Spain did rise substantially due to rising risk aversion.

All these effects should be taken into account by professional analysts. Hence, for political decision makers forecasts provided by professional analysts might be a welcome aid during crisis events. However, following our empirical findings this conclusion might be somewhat overhasty. We found strong evidence for rising forecast errors (at least for the survey mean) and rising disagreement amongst forecasters during financial crisis events. This puts the usefulness of survey forecasts into question. As a matter of fact, blindly following professional analysts' predictions might lead to wrong decisions. By not focusing solely on the accuracy of the forecasts but also on the implications of uncertainty measures we are able to get crisis related insights from professional analysts' predictions and are able to link the uncertainty measures to the real economy. Especially the immediate rise in the HLS for the various investigated forecasts might serve political decision makers as a discretional warning sign. Hence, the ongoing monitoring of uncertainty measures might be much more useful than focusing solely on point forecasts.

8. Conclusion

In this paper the forecasts of professional analysts for the three months interbank rate as well as the 10Y government bond yield in the United Kingdom and Germany have been evaluated. Applying basic measures of forecast accuracy it has been shown, that the UK survey forecasts in general do not perform better than a simple naive forecast. These results do fit to the findings of earlier studies dealing with survey forecasts for German interest rates. Additionally, evaluating the mean forecasts for the United Kingdom and Germany using models of time series analysis does not lead to the conclusion that the survey forecasts deliver useful information (with the 10Y4M survey mean for UK as the only exception) with regard to forecast accuracy.

This does not necessary has to lead to the conclusion that survey forecasts are useless in general. Because of that, we have further investigated forecast dispersion measures. As a matter of fact, in financial crisis uncertainty is a major concern central bank officials, political decision makers, and financial market participants. Applying our empirical approach we were able to find crisis related implications and we were able to show, that mean shifts in the uncertainty measures coincide with times of financial turmoil.

This raises the question whether uncertainty measures might serve as some kind of warning signal or early warning indicator. Using a simple approach we found empirical evidence for causal relationships between the uncertainty measures and changes in real economic indicators for UK and Germany (industrial production). But more research has to be done. Especially due to omitting variables and/or long memory in the time series more robustness checks are inevitably necessary. Furthermore uncertainty measures for a broader range of economic and financial variables should be taken into account.

References

Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets, or surveys forecast inflation better? J. Monet. Econ. 54, 1163–1212.
Baghestani, H., 2006. An evaluation of the professional forecasts of US long-term interest rates. Rev. Finan. Econ. 15, 177–191.
Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. J. Appl. Econometr. 18, 1–22.
Basse, T., Reddemann, S., 2010. Variable-ordering induced problems of impulse-response analysis and other difficulties: the dividend policy of Austrian firms. Int. J. Comput. Econ. Econometr. 1, 278–293.
Belongia, M., 1987. Predicting interest rates: a comparison of professional and market-based forecasts. In: Federal Reserve Bank of St. Louis Review, pp. 9–15.
Bernanke, B., 1990. On the predictive power of interest rates and interest rate spreads. Tech. rep. National Bureau of Economic Research.
Bernanke, B.S., Blinder, A.S., 1992. The federal funds rate and the channels of monetary transmission. In: The American Economic Review, pp. 901–921.
Bernard, H., Gerlach, S., 1998. Does the term structure predict recessions? The international evidence. Int. J. Finan. Econ. 3, 195–215.
Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77, 623–685.
Bomberger, W.A., 1996. Disagreement as a measure of uncertainty. J. Money Credit Bank. 28, 381–392.
Cheung, Y.-W., Chinn, M.D., 1998. Integration, cointegration and the forecast consistency of structural exchange rate models. J. Int. Money Finan. 17, 813–830.
Chortareas, G., Jitmaneeroj, B., Wood, A., 2012. Forecast rationality and monetary policy frameworks: evidence from UK interest rate forecasts. J. Int. Finan. Markets Inst. Money 22, 209–231.
Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 74, 427–431.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Business Econ. Stat. 13, 253–263.
Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. CRC Press.
Estrella, A., Rodrigues, A.P., Schich, S., 2003. How stable is the predictive power of the yield curve? Evidence from Germany and the United States. Rev. Econ. Stat. 85, 629–644.
Friedman, B.M., 1980. Survey evidence on the rationality of interest rate expectations. J. Monet. Econ. 6, 453–465.
Givoly, D., Lakonishok, J., 1984. The quality of analysts' forecasts of earnings. Finan. Anal. J. 40, 40–47.
Granger, C.W., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: J. Econometr. Soc., 424–438
Greer, M., 2003. Directional accuracy tests of long-term interest rate forecasts. Int. J. Forecast. 19, 291–298.
Gupta, R., Hammoudeh, S., Modise, M.P., Nguyen, D.K., 2014. Can economic uncertainty, financial stress and consumer sentiments predict US equity premium? J. Int. Finan. Markets Inst. Money 33, 367–378.
Hafer, R.W., Hein, S.E., 1989. Comparing futures and survey forecasts of near-term treasury-bill rates. Federal Reserve Bank St. Louis Rev. 71, 33–42.
Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22, 679–688.
Ince, O., Molodtsova, T., 2017. Rationality and forecasting accuracy of exchange rate expectations: evidence from survey-based forecasts. J. Int. Finan. Markets Inst. Money 47, 131–151.
Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica: J. Econometr. Soc., 1551–1580.
Jongen, R., Verschoor, W.F., 2008. Further evidence on the rationality of interest rate expectations. J. Int. Finan. Markets Inst. Money 18, 438–448.
Kolb, R., Stekler, H.O., 1996. How well do analysts forecast interest rates? J. Forecast. 15, 385–394.
Kunze, F., Kramer, J., Rudschuck, N., 2014. Interest rate forecasts in times of financial crisis: what might be interesting to know? Eur. J. Polit. Econ. 34, S45–S52.
Lahiri, K., Sheng, X., 2010. Measuring forecast uncertainty by disagreement: the missing link. J. Appl. Econometr. 25, 514–538.
Miah, F., Khalifa, A.A., Hammoudeh, S., 2016. Further evidence on the rationality of interest rate expectations: a comprehensive study of developed and emerging economies. Econ. Modell. 54, 574–590.
Mitchell, K., Pearce, D.K., 2007. Professional forecasts of interest rates and exchange rates: evidence from the wall street journals panel of economists. J. Macroecon. 29, 840–854.
Pesaran, M.H., Weale, M., 2006. Survey Expectations. Handbook of Economic Forecasting, vol. 1, pp. 715–776.
Pfaff, B., 2008. Analysis of Integrated and Cointegrated Time Series with R. Springer Science & Business Media.
Phillips, P.C., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.
Ploberger, W., Krämer, W., Kontrus, K., 1989. A new test for structural stability in the linear regression model. J. Econometr. 40, 307–318.
Rich, R.W., Raymond, J.E., Butler, J.S., 1992. The relationship between forecast dispersion and forecast uncertainty: evidence from a survey data-arch model. J. Appl. Econometr. 7, 131–148.
Saar, D., Yagil, Y., 2015. Forecasting growth and stock performance using government and corporate yield curves: evidence from the European and Asian markets. J. Int. Finan. Markets Inst. Money 37, 27–41.

Schmeling, M., Schrimpf, A., 2011. Expected inflation, expected stock returns, and money illusion: what can we learn from survey expectations? Eur. Econ. Rev. 55, 702–719.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.
Schwarzbach, C., Kunze, F., Rudschuck, N., Windels, T., 2012. Asset management in the german insurance industry: the quality of interest rate forecasts. German J. Risk Insurance 101, 693–703.
Sibbertsen, P., Wegener, C., Basse, T., 2014. Testing for a break in the persistence in yield spreads of EMU government bonds. J. Bank. Finan. 41, 109–118.
Simon, D.P., 1989. The rationality of federal funds rate expectations: evidence from a survey. J. Money Credit Bank. 21, 388–393.
Sims, C.A., 1980. Macroeconomics and reality. Econometrica: J. Economet. Soc., 1–48
Spiwoks, M., Bedke, N., Hein, O., 2008. Forecasting the past: the case of US interest rate forecasts. Finan. Markets Portfolio Manage. 22, 357–379.
Stillwagon, J.R., 2015. Testing the expectations hypothesis with survey forecasts: the impacts of consumer sentiment and the zero lower bound in an I(2) CVAR. J. Int. Finan. Markets Inst. Money 35, 85–101.
Stock, J.H., Watson, M.W., 1989. Interpreting the evidence on money-income causality. J. Econometr. 40, 161–181.
Theil, H., 1992. Who forecasts best? In: Henri Theils Contributions to Economics and Econometrics. Springer, pp. 1115–1120.
Wegener, C., Spreckelsen, C., Basse, T., Mettenheim, H.-J., 2016. Forecasting government bond yields with neural networks considering cointegration. J. Forecast. 35, 86–92.
Zarnowitz, V., 1992. Consensus and uncertainty in economic prediction. In: Business Cycles: Theory, History, Indicators, and Forecasting. University of Chicago Press, pp. 492–518.
Zeileis, A., 2006. Implementing a class of structural change tests: an econometric computing approach. Comput. Stat. Data Anal. 50, 2987–3008.
Zeileis, A., Kleiber, C., Krämer, W., Hornik, K., 2003. Testing and dating of structural changes in practice. Comput. Stat. Data Anal. 44, 109–123.

Chapter 4

Oil prices and sovereign credit risk of oil producing countries: an empirical investigation

with Christoph Wegener, Tobias Basse and Hans-Jörg von Mettenheim
contribution Frederik Kunze: 20%

Published:
Quantitative Finance, Vol. 16 Issue: 12, pp. 1961–1968,
https://doi.org/10.1080/14697688.2016.1211801

Abstract

The low oil price recently has caused fears about the sustainability of public finances in some oil producing countries. We examine the relationship between oil prices and sovereign credit risk examining the CDS market. Analysing data from nine countries (Brazil, Malaysia, Norway, Qatar, Russia, Saudi Arabia, the United Kingdom, the United States of America and Venezuela) we have estimated bivariate VAR-GARCH-in-mean models. The results of our empirical investigations generally speaking do suggest that positive oil price shocks lead to lower sovereign CDS spreads. Thus, our findings support the hypothesis that higher oil prices improve the fiscal stability of oil producing countries.

Keywords

Sovereign debt crisis, oil prices, fiscal stability.

JEL Classifications

G17, H63, C58, Q43

Highlights

- We investigate daily observations of credit default swap (CDS) spreads of eight oil producing countries and the price of crude oil.
- We employ VAR(1)-GARCH(1,1)-in-mean models to draw inference about the impact of oil price changes on the market perception of sovereign credit risk.
- We document the inverse relationship between oil prices and sovereign CDS spreads for Saudi Arabia, Brazil, Malaysia, Russia, Qatar and Venezuela.
- For Norway, the United Kingdom and the U.S. we did not find a significant inverse relationship and attribute these results to the diversified character of these economies.
- Summarizing our results, it can be stated that crude oil prices seem to matter for the market perception of sovereign credit risk and fiscal stability in oil producing countries.
- Given the relevance of sovereign credit risk this results might be of high impact for financial market practitioners and political decision-makers.

Chapter 5

The usefulness of oil price forecasts – evidence from survey predictions

with Markus Spiwoks, Kilian Bizer and Torsten Windels
contribution Frederik Kunze: 70%

Published:
Managerial and Decision Economics, Vol. 39 Issue: 4, pp. 427–446,
http://dx.doi.org/10.1002/mde.2916

Abstract

This paper evaluates survey forecasts for crude oil prices and discusses the implications for decision makers. A novel disaggregated data set incorporating individual forecasts for Brent and WTI is used. We carry out tests for unbiasedness, sign accuracy and forecast encompassing, followed by the computation of coefficients for topically-oriented trend adjustments and the Theil's U measure. We also control for the forecast horizon finding heterogeneous results. Forecasts are more precise for shorter horizons, but less accurate than the naïve prediction. For longer horizons, topically-oriented trend adjustments become more pronounced, but forecasters tend to outperform the naïve predictions.

Keywords

Criteria for decision-making under risk and uncertainty, forecasting and simulation: models and applications, international finance forecasting and simulation: models and applications

JEL Classifications

D81, F37, F47

Highlights

- We utilize a novel comprehensive disaggregated dataset incorporating individual forecasts for both Brent and WTI crude oil prices.
- We calculate prominent measures of forecast accuracy and carry out tests for unbiasedness, sign accuracy and forecast encompassing to assess the rationality and accuracy of professional forecasters.
- We discuss implications for decision-makers as regards the applications of oil price forecasts.
- Forecasts are more precise for shorter horizons, but less accurate than the naïve prediction.
- For longer horizons, topically-oriented trend adjustments become more pronounced, but forecasters tend to outperform the naïve predictions.

Received: 15 May 2017 | Revised: 26 October 2017 | Accepted: 15 January 2018 82
DOI: 10.1002/mde.2916

RESEARCH ARTICLE

WILEY

The usefulness of oil price forecasts—Evidence from survey predictions

Frederik Kunze[1,3] | Markus Spiwoks[1,2] | Kilian Bizer[1] | Torsten Windels[3]

[1] Faculty of Economic Sciences, Chair of Economic Policy and SME Research, Georg-August-University, Göttingen, Germany

[2] Faculty of Business Administration, Ostfalia University of Applied Sciences, Wolfsburg, Germany

[3] NORD/LB Norddeutsche Landesbank Girozentrale, Hannover, Germany

Correspondence
Frederik Kunze, NORD/LB Norddeutsche Landesbank Girozentrale, Hannover, Germany.
Email: frederik.kunze@nordlb.de

JEL Classification: D81; F37; F47

This paper evaluates survey forecasts for crude oil prices and discusses the implications for decision makers. A novel disaggregated data set incorporating individual forecasts for Brent and Western Texas Intermediate is used. We carry out tests for unbiasedness, sign accuracy, and forecast encompassing, followed by the computation of coefficients for topically oriented trend adjustments and the Theil's U measure. We also control for the forecast horizon finding heterogeneous results. Forecasts are more precise for shorter horizons, but less accurate than the naïve prediction. For longer horizons, topically oriented trend adjustments become more pronounced, but forecasters tend to outperform the naïve predictions.

1 | INTRODUCTION

Crude oil prices have a direct influence on global economic activity. In addition, the recent low price environment calls into question the sustainability of public finances in a rising number of oil exporting countries (see, e.g., Wegener, Basse, Kunze, & von Mettenheim, 2016). For monetary policy-makers at major central banks, the outlook regarding the price of crude oil holds vital importance for their inflation and growth forecasts (see, e.g., Knetsch, 2007; as well as Natal, 2012). Oil prices as an input factor are also important for strategical decision making and corporate profitability (see, e.g., Sadorsky, 2008). For example, Haushalter, Heron, and Lie (2002) have emphasized the impact of oil price uncertainty on corporate value. Furthermore, crude oil prices have an impact on stock markets both in emerging and developed markets (see, e.g., Basher & Sadorsky, 2006; as well as Miller & Ratti, 2009). Given the risks and uncertainties related to global oil prices and, hence, transaction costs for corporations, the future expectations regarding crude oil prices and accurate oil price predictions are essential in terms of understanding the outlook for the global economy and for individual firms.

As regards, for example, information systems, forecasts belong to the methodologies of decision support systems (see, e.g., Banker & Kauffman, 2004) and, hence, good predictions are essential for efficient respectively effective management decisions and planning processes (see Waddell & Sohal, 1994). Accordingly, the necessity for forecasts is a result of the general uncertainty firms have to deal with (see Lowe, 1970).

Various forecasting models and the usefulness of corresponding forecasts have been empirically assessed (see, e.g., Alquist, Kilian, & Vigfusson, 2013). Regarding the applied forecasting methodologies for crude oil prices, Behmiri and Pires Manso (2013) and Frey, Manera, Markandya, and Scarpa (2009) provide extensive overviews. Comparing the framework of global oil markets with "purely" financial markets (e.g., bond, equity, and FX markets), one might conclude that oil prices should be somewhat more straightforward to forecast due to the available fundamental market information regarding consumption patterns and physical supply (i.e., global crude oil production and crude oil reserves). Within this line of thought, the global balance between supply and demand and the respective expectations should at least be suitable to forecast the long-term price of oil. However, for corporate managers, long-term outlooks might not be adequate for decisions with a short- or medium-term impact. In addition, forecasters have to account for the endogenous relationship between oil prices and economic activity (see He, Shouyang, & Lai, 2010). Furthermore, shock events on both demand and supply side may lead to a higher forecast risk despite the available information. Supply side shocks have inter alia been investigated by Kilian (2008) as well as Kilian (2009). Kilian (2009) mentioned three types of shocks: oil supply shocks, aggregate demand shocks, and precautionary demand shocks. Because the OPEC countries are major suppliers of crude oil, there exists extensive literature investigating the price impact of the OPEC meetings and the official announcements of the cartel members (see, e.g., Bentzen, 2007; Kaufmann, Dees, Karadeloglou, & Sanchez, 2004; Schmidbauer & Rösch, 2012). Besides these fundamental effects, speculation also plays a role when discussing the usefulness of oil price forecasts (see Kaufmann & Ullman, 2009; Prat & Uctum, 2011).

Given the various factors influencing crude oil prices, the accuracy of oil price forecasts is important for decision-makers. Hence, this study aims to evaluate survey predictions for the price of Brent and Western Texas Intermediate (WTI). The remainder of the paper is structured as follows. In Section 3, we will present the dataset, before describing the evaluation methods employed in Section 4. Subsequently, in Section 5, we present the empirical findings, and finally, Section 6 concludes the paper.

2 | SURVEY PREDICTIONS FOR THE PRICE OF OIL

As mentioned previously, a wide range of forecasting approaches for oil prices exists. Decision-makers and managers have to choose among them and have to decide how they apply these forecasts as a decision support tool. Given numerous influences and potential shock effects, results from survey predictions may be seen as one alternative (see Alquist et al., 2013) for corporate managers. Survey forecasts for oil prices respectively financial market variables in general are useful in several ways for decision-makers. For example, Pesaran and Weale (2006) have noted that “combinations of forecasts produced by different bodies tend to be more accurate than forecasts produced by any individual” (see Pesaran & Weale, 2006, p. 748). Furthermore, corporations not able or willing to produce own predictions have recourse to regularly produced predictions. Rötheli (1998) has investigated the rivalry between forecasting and other strategies (i.e., diversification or waiting) and came inter alia to the conclusion that high forecasting costs lead to a preference of alternative strategies. Hence, survey forecasts when available at comparatively low costs might constitute an alternative. Decision-makers from institutions that are producing their own forecasts have a helpful tool when benchmarking the own predictions. Hence, forecasts may create economic value. However, given the various factors influencing economic variables, a perfect forecast is far from certain. Hence, it is essential for decision-makers to appraise the forecasts and to develop an understanding of the errors in the past and the track record of available forecasting methods or forecasters (see, e.g., Barron & Targett, 1988).

Despite this global importance, to the best of our knowledge, the academic literature dealing with the evaluation of survey forecasts for the price of crude oil is rather scarce, at least in comparison with the assessment of other financial market variables such as interest rates, corporate earnings, or exchange rates. There exists a large body of literature dealing with the evaluation of survey predictions for exchange rate forecasts (see, e.g., Audretsch & Stadtmann, 2005; Dominguez, 1986; Mitchell & Pearce,2007; Pierdzioch & Rülke, 2015; Takagi, 1991; and more recently Ince & Molodtsova, 2017) and interest rate forecasts (see, e.g., Belongia, 1987; Benke, 2006; Friedman, 1980; Kolb & Stekler, 1996; Mitchell & Pearce, 2006; Spiwoks, Bedke, & Hein, 2008; Spiwoks, Bedke, & Hein, 2010; only to name but a few). Recently, for example, Reitz, Stadtmann, and Taylor (2010), Jongen, Verschoor, Wolff, and Zwinkels (2012), and Beckmann and Czudaj (2017) focus on FX forecasters' heterogeneity. Swanson (2006) investigates inter alia the dispersion of Blue Chip Forecasts for the 3 months' Treasury Bill Rate.

One possible explanation for the relative scarcity of studies dealing with survey forecasts for oil prices is the bounded availability of surveys collecting oil price forecasts. The following sources for survey forecasts are mentioned in the relevant empirical literature. On a monthly basis, survey predictions for the price of oil are collected by Consensus Economics. MacDonald and Marsh (1993) conclude that individual forecasts are irrational (i.e., are biased and inefficient), whereas Reitz, Rülke, and Stadtmann (2009) find evidence for the underperformance of forecasters compared with a random walk forecast. Prat and Uctum (2011) reject the hypothesis of rational expectations regarding forecasts for the price of WTI. For the sake of completeness, one could also refer to Singleton (2013), who inter alia investigates the dispersion of oil price forecasts collected by Consensus Economics. Pierdzioch, Rülke, and Stadtmann (2010), Pierdzioch, Rülke, and Stadtmann (2013), and Rülke et al. (2012) evaluate the oil price forecasts from the ECB's survey of professional forecasters (SPF). Pierdzioch et al. (2010) find evidence for antiherding of professional forecasters, whereas Rülke et al. (2012) conclude that forecasters do not act rationally, at least in terms of internal consistency regarding the forecast horizon. More recently, Atalla, Joutz, and Pierru (2016) examined the dispersion of individual forecasts from the ECB's SPF. Within this paper, we evaluate the survey forecasts collected by the news agency Dow Jones International GmbH. The relevance of forecast evaluation in the context of managerial decision-making has been inter alia discussed by Lawrence (1991).

To the best of our knowledge, our study is the first to assess the quality of forecasts from the Dow Jones Oil Price Survey. With this new dataset, as regards academic research, we are able to deliver novel empirical evidence in terms of forecast evaluation for the price of oil. In addition to widely acknowledged evaluation measures (e.g., mean average error or root mean squared error, RMSE), we also perform tests for sign accuracy, unbiasedness, and forecast encompassing.

3 | THE DATASET

We examine survey data collected by the news agency Dow Jones International GmbH. The use of survey forecasts for economic and financial market variables inter alia allows to test for the rationality of the professional analysts (see, e.g., MacDonald, 1990). In the case of the Dow Jones Oil price survey, professional financial analysts as well as researchers from economic think tanks are requested to provide their forecasts for benchmark energy prices (i.e., the price for the sweet light crude oil from the North Sea Brent and the lighter U. S. Benchmark WTI). In general, we denote the actual price of Brent as $p_{Brent,\ i}$, whereas $p_{WTI,\ i}$ is the actual price of WTI. Furthermore, the corresponding forecasts are denoted as $\hat{p}^{j}_{Brent,i}$ and $\hat{p}^{k}_{WTI,i}$, respectively.

The monthly data ranges from January 2006 to December 2015, and hence incorporates substantial movements in the price of oil. Theoretically, 468 point forecasts for Brent and WTI could have been delivered by every single forecast institution. One important issue of the dataset is the substantial number of missing values (NAs). This leads to a strongly unbalanced dataset. The raw dataset contains forecasts from 75 individual providers and sums up to 22,759 (10,180 WTI and 12,579 Brent forecasts) individual predictions.[1] To ensure a higher level of statistical validity, forecasters with fewer than 40 predictions for the price of Brent and WTI, have been excluded in a first step. This adjustment leads to a dataset with 21,907 individual forecasts. In the time period under investigation, the Brent price forecasts of j = 52 institutions and the WTI price forecasts of k = 50 institutions are investigated. Additionally, for every forecast i for Brent or WTI, a survey median is reported. The naïve forecast is given by the actual forecast at the date of publication of the survey (forecast date). It is important to note that due to the specific process of data collection, the naïve forecast for each forecast horizon is not a shift of the actual oil price.

In addition, some further prior considerations have to be made due to the structure of the survey. Every month on the forecast date, the forecasters are requested to deliver the forecast for the end of quarter prices for Brent (and WTI) for the following four quarters. For forecast dates in the last month of the current quarter, the requested forecast horizon shifts by one quarter (see also Table 1 below).

Following the sample above, on February 17, 2010, the final forecast for the March 31, 2010 end of quarter forecast had to be delivered. On March 17, 2010, the first forecast for March 31, 2011 had to be delivered. Due to the monthly frequency of the data for a single forecast date, there exist up to 12 forecasts, and hence up to 12 forecast horizons. Owing to calendar effects and minor variations in the forecast date, the forecast horizons slightly change over time. Table 2 below shows the corresponding forecast horizons. The structure of the data described above leads to substantial overlapping, which has to be accounted for. Accordingly, we start our empirical analysis in Section 5 with evaluation methods that can be applied without consideration of the forecast horizon.

In addition to the forecasts, the data set contains the actual forecast (i.e., the price of oil at the end of quarter date) and the naïve forecast (i.e., the price of oil at the forecast date). Given the specific characteristics of the Dow Jones Oil Price Survey, the dataset has some clear advantages. First, with up to 52 contributors (after adjustments), the group of forecasters is rather large in comparison with other existing surveys. This fact is especially useful for decision makers, because the disaggregated data set allows to analyze the range of the survey forecasts for the oil price (see, e.g., Gripaios, 1994). In general, survey forecasts allow to test hypothesis in the context of uncertainty of financial market variables. In this context, Lahiri and Sheng (2010), Atalla et al. (2016), and Glas and Hartmann (2016) investigated disaggregated data sets to gain insights from the dispersion of individual forecasts. Atalla et al. (2016) came inter alia to the conclusion that dispersion of oil price forecasts is positively correlated with forecast errors. Following Atalla et al. (2016), who investigated the ECB's SPF oil price predictions, an increased disagreement hence indicates higher uncertainty. Lahiri and Sheng (2010) empirically assessed inflation and output forecasts contributed to the Federal Reserve Bank of Philadelphia's SPF and came to the conclusion that disagreement is a fairly good measure of uncertainty in low volatility periods and for shorter forecast horizons, respectively. A second advantage of the data used in this paper stems from the fact that two globally recognized crude oil benchmark forecasts are requested. In fact, a substantially large fraction of the analysts deliver forecasts for both Brent and WTI. Furthermore, the existence of no fewer than 12 forecast horizons allows drawing conclusions regarding the forecast horizon. Finally, the dataset covers the global financial crisis, the following rebound in global economic activity and the price collapse attributed to the shale oil boom in the United States.

To illustrate the underlying data, in Figures 1 and 2, the actual Brent price, the median forecast, and the naïve forecast are shown, whereas Figures 3 and 4 show the corresponding time series for WTI. Regarding the naïve prediction, it should be considered that —as noted above—the naïve prediction is given by the price of oil at the forecast date. Hence, the naïve forecast is not a simple shift of the actual oil price and, thus, it does not perfectly run after the actual oil price.

With a forecast horizon of 1 month (Figure 1 for Brent and Figure 3 for WTI), the forecast error of the median forecast is much smaller than in the case of the 12-month horizon (Figure 2 for Brent and Figure 4 for WTI). Furthermore, it can easily be seen that the naïve

TABLE 2 Forecast horizons

Forecast horizon	Number of weeks (rounded)
H1	6
H2	10
H3	15
H4	19
H5	23
H6	28
H7	32
H8	36
H9	40
H10	45
H11	50
H12	54

TABLE 1 Structure of the data set

Forecast date	Forecast end of quarter	Actual price (USD/barrel)	Naïve forecast (USD/barrel)	Individual forecast (USD/barrel)
February 17, 2010	31.03.2010	81.0	76.0	74.0
	30.06.2010	74.0	76.0	75.0
	30.09.2010	81.0	76.0	78.3
	31.12.2010	94.0	76.0	80.0
March 17, 2010	30.06.2010	74.0	81.0	75.0
	30.09.2010	81.0	81.0	76.0
	31.12.2010	94.0	81.0	80.0
	31.03.2011	117.0	81.0	81.0

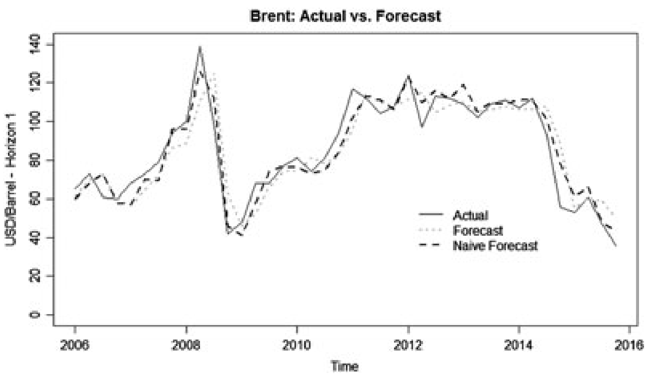


FIGURE 1 Crude oil price Brent (actual vs. forecast)—Forecast horizon 1 month

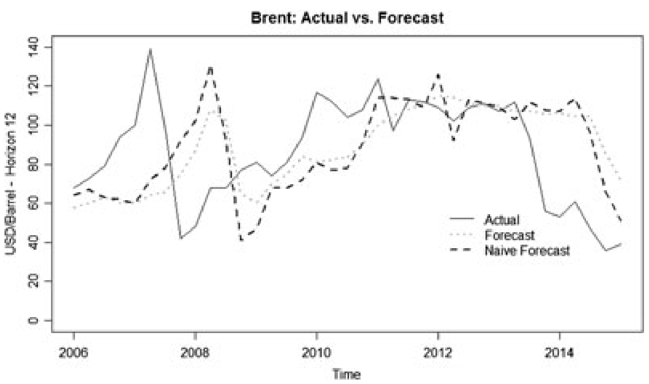


FIGURE 2 Crude oil price Brent (actual vs. forecast)—Forecast horizon 12 months

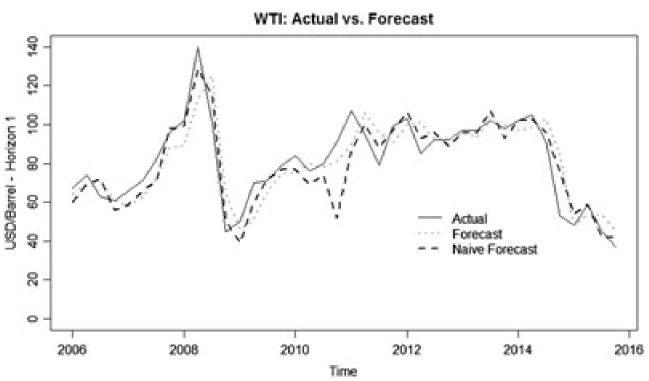


FIGURE 3 Crude oil price Western Texas Intermediate (WTI; actual vs. forecast)—Forecast horizon 1 month

prediction and the median forecast lie close together for both horizons, which is an indication for a status quo bias (i.e., topically oriented trend adjustment) of crude oil forecasters.

4 | METHODOLOGIES OF FORECAST EVALUATION

Following a standard approach in the context of forecast evaluation (see, e.g., Beechey & Österholm, 2014; as well as Baghestani, Arzaghi, & Kaya, 2015), we start our empirical analysis with the test for unbiasedness, followed by the sign accuracy test and the test for forecast encompassing. In a second step, we apply widely acknowledged forecast accuracy measures.

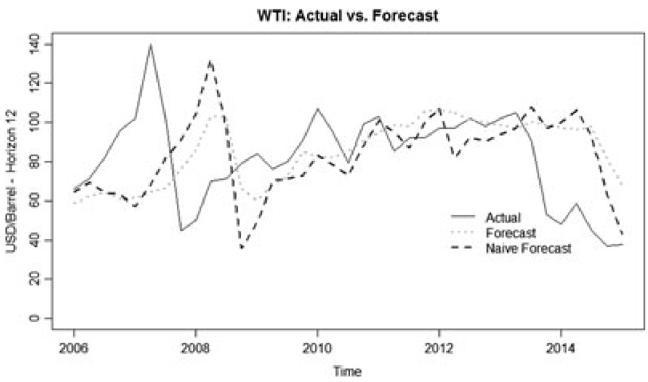


FIGURE 4 Crude oil price Western Texas Intermediate (WTI; actual vs. forecast)—Forecast horizon 12 months

4.1 | Test for unbiasedness

In the context of forecast evaluating, the test for unbiasedness is a common test procedure to check for systematic (i.e., nonrandom) errors (see, e.g., Abosedra & Baghestani, 2004; Friedman, 1980; Holden & Peel, 1990; Mincer & Zarnowitz, 1969; and more recently Bowles et al., 2010; Kenny, Kostka, & Masera, 2015; Pierdzioch et al., 2013; as well as Spiwoks, Scheier, & Hein, 2014).

We test for unbiased forecasts for the change in the price of Brent, with $\Delta p_{Brent,\ i}$ being the actual price change and $\Delta\widehat{p}^{j}_{Brent,i}$ being the forecast for the change in the price of Brent from institution *j*. The actual change of the price of one barrel WTI is given by $\Delta p_{WTI,\ i}$, and the forecasted price change is denoted with $\Delta\widehat{p}^{j}_{WTI,i}$. In Formula 1 below, Brent Model 1 and WTI Model 1 are unbiased forecasts. Hence, to statistically validate the absence of a systematic forecast error in Brent Model 2 and WTI Model 2, α has to be equal to 0 and β has to be equal to 1. As discussed earlier, the data set contains forecasts for up to 12 forecast horizons. However, we begin by applying the test for unbiasedness regardless of the forecast horizon.

$$
\begin{aligned}
&\text{Unbiased forecast (Brent Model 1): } \Delta p_{Brent,i} = \Delta\hat{p}^{\,j}_{Brent,i} + u^j_i \\
&\text{Biased forecast (Brent Model 2): } \Delta p_{Brent,i} = \alpha + \beta\Delta\hat{p}^{\,j}_{Brent,i} + u^j_i \\
&\text{and} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (1) \\
&\text{Unbiased forecast (WTI Model 1): } \Delta p_{WTI,i} = \Delta\hat{p}^{\,k}_{WTI,i} + u^k_i \\
&\text{Biased forecast (WTI Model 2): } \Delta p_{WTI,i} = \alpha + \beta\Delta\hat{p}^{\,j}_{WTI,i} + u^k_i .
\end{aligned}
$$

4.2 | Sign accuracy test

By applying the sign accuracy test, we investigate the directional accuracy of the forecasts provided (see, e.g., Baghestani, 2008; Baghestani et al., 2015; Diebold & Lopez, 1996; Greer, 2003; Kolb & Stekler, 1996; Mitchell & Pearce, 2006; as well as Spiwoks, Bedke, & Hein, 2009). Following—for example—Spiwoks et al. (2009), we first use a 2-by-2 contingency table. In the fields N_{11}

and N_{22} of Table 3 below, the correct forecasts for the directional change of the oil price can be found. Incorrect directional forecasts for the price of oil can be found in the fields N_{21} and N_{12}. As with the test for unbiasedness, we do not differentiate for the forecast horizon.

Following Diebold and Lopez (1996), we test for significant differences between a random directional forecast and the crude oil forecasts provided in the data set using a χ^2 test. Whenever this is the case, we test for superiority of the forecast over the random directional change.

4.3 | Test for forecast encompassing

Furthermore, we want to check for additional information content of the forecasts against a simple benchmark model using the test for forecast encompassing (see, e.g., Diebold & Lopez, 1996; as well as Spiwoks et al., 2009; Clements and Harvey, 2010). In a first step, we apply the setup proposed by McCracken and West (2008) and estimate a simple linear regression model for the actual price, the naïve prediction (as the simple benchmark forecast), and the delivered survey forecasts (see Formula 2 below). Under the null hypothesis (H_0), the benchmark model encompasses the survey forecast, and hence β_1 should be 1 and β_2 should be 0.

$$p_{Brent,i} = \alpha + \beta_1 \hat{p}_{Brent}^{Naive} + \beta_2 \hat{p}^{\,j}_{Brent} + u_i \quad \text{and} \quad p_{WTI,i} = \alpha + \beta_1 \hat{p}_{WTI}^{Naive} + \beta_2 \hat{p}_{WTI}^{k} + u_i. \tag{2}$$

As a robustness check for the results, we will also apply these test procedures while controlling for the forecast horizon. Due to the challenges regarding the large number of missing values within the data, we only use the median forecast when calculating the forecast horizon-dependent results, as this is the only forecast time series without gaps.

Following the test procedures outlined above, we assess the forecasts with common forecast accuracy measures: the mean of the forecast errors, the (adjusted) RMSE, the TOTA coefficient, and the Theil's U score. One clear advantage of these measures stems from the fact that we are able to directly compare different forecasts in terms of their accuracy. Once again, in a first step, we do not control for the forecast horizon. In a second step, to derive implications regarding forecast accuracy depending on the forecast horizon, we additionally investigate the measures of forecast accuracy controlling for the forecast horizon at the level of individual forecasts.

TABLE 3 Sign accuracy test 2-by-2 contingency table

	Actual event: Oil price rises	Actual event: Oil price falls	Σ
Forecast: Oil price rises	N_{11}	N_{12}	$N_{1.}$
Forecast: Oil price falls	N_{21}	N_{22}	$N_{2.}$
Σ	$N_{.1}$	$N_{.2}$	N

4.4 | Forecast error measures

We first derive the mean forecast error as the weighted difference between the actual price for Brent and WTI (see Formula 3 below). One advantage of this accuracy measure stems from the fact that it is easy to both calculate and interpret. For example, a positive value of the mean forecast error measure for forecaster *j* indicates that *j* overestimates the price of oil on average. A comprehensive overview of forecast evaluation measures—including describing the mean forecast error—is provided by Andres and Spiwoks (1999).

$$\bar{e}^{j}_{Brent,h} = \frac{1}{N}\sum_{i=1}^{N}\left(p_{Brent,i,h} - \hat{p}^{\,j}_{Brent,i,h}\right) \quad \text{and} \quad \bar{e}^{k}_{WTI,h} = \frac{1}{N}\sum_{i=1}^{N}\left(p_{WTI,i,h} - \hat{p}^{k}_{WTI,i,h}\right). \tag{3}$$

The forecast errors for the price of Brent for the forecaster *j* are denoted as e^{j}_{Brent}, and the forecast errors for the price of WTI for the forecaster *k* are denoted as e^{k}_{WTI}. Once again, $p_{Brent,\,i}$ is the actual price of Brent, whereas $p_{WTI,\,i}$ is the actual price of WTI. Although numerous forecasters deliver forecasts for both Brent and WTI, it does not have to be the case necessarily that *j* = *k*. The corresponding forecasts are denoted as $\hat{p}^{j}_{Brent,i}$ and $\hat{p}^{k}_{WTI,i}$, respectively. The relevant forecast horizon with *h* = 1, 2, …, 12 is denoted by *h*.

Aside from the easy-to-interpret mean forecast error, quadratic error measures are widely used within the evaluation literature. One advantage stems from the fact that positive and negative forecast errors do not level out. Furthermore, using a quadratic error measure leads to a penalization of comparatively large forecast errors. The RMSE given in Formula 4 below is one very commonly applied error measure (see, e.g., Andres & Spiwoks, 1999; Espinoza, Fornari, & Lombardi, 2012; Hyndman & Koehler, 2006; Kisinbay, 2010; Leitch & Tanner, 1991; and more recently Herwartz & Schlüter, 2017; Ryan & Whiting, 2017; as well as Wegener, Spreckelsen, Basse, & Mettenheim, 2016).

$$RMSE^{j}_{Brent,h} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(p_{Brent,i,h} - \hat{p}^{j}_{Brent,i,h}\right)^2} \quad \text{and} \quad RMSE^{k}_{WTI,h} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(p_{WTI,i,h} - \hat{p}^{k}_{WTI,i,h}\right)^2}. \tag{4}$$

As stated above, the RMSE is widely used in the relevant literature dealing with accuracy of forecasts. However, following—for example—Clements (2014) and Chen, Costantini, and Deschamps (2016) under certain circumstances, it may be appropriate to apply an adjusted version of the RMSE (aRMSE). Given the large movements in the oil price and the high likelihood of large forecast errors in the context of the oil price crisis, individual forecasters who missed out sequences of high uncertainty might be “rewarded” by the conventional RMSE measure. Against the background of the unbalanced panel data, there is a high likelihood for such a misspecification.

The aRMSE given in Formula 5 below attributes a lower weight to forecast errors when the cross-sectional forecast error is high. In the context of our research questions, the aRMSE measure seems to be more appropriate. Following Clements (2014) and Chen et al. (2016), we use the median of the absolute forecast errors rather than the mean forecast errors.

$$aRMSE^{j}_{Brent,h} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(p_{Brent,i,h} - \hat{p}^{j}_{Brent,i,h}\right)^{2}}\,\frac{median_i\left(median_j\left(\left|p_{Brent,i,h} - \hat{p}^{j}_{Brent,i,h}\right|\right)\right)}{median_j\left(\left|p_{Brent,i,h} - \hat{p}^{j}_{Brent,i,h}\right|\right)}.$$

and (5)

$$aRMSE^{k}_{WTI,h} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(p_{WTI,i,h} - \hat{p}^{k}_{WTI,i,h}\right)^{2}}\,\frac{median_i\left(median_k\left(\left|p_{WTI,i,h} - \hat{p}^{k}_{WTI,i,h}\right|\right)\right)}{median_k\left(\left|p_{WTI,i,h} - \hat{p}^{k}_{WTI,i,h}\right|\right)}.$$

4.5 | Topically oriented trend adjustment

We will also check for topically oriented trend adjustment using the TOTA coefficient developed by Andres and Spiwoks (1999). The TOTA coefficient has been especially used in the context of evaluating interest rate forecasts (Spiwoks et al., 2008; Spiwoks et al., 2009; Kunze, 2014 as well as Kunze et al. 2015), economic forecasts (see, e.g., Spiwoks et al., 2014), and FX forecasts (see, e.g., Bofinger & Schmidt, 2003; as well as Baghestani, 2010b). Meub, Proeger, Bizer, and Spiwoks (2015) have also used the TOTA coefficient in an experimental context. The TOTA coefficient is shown in Formula 6 below. In the context of our investigation, the TOTA coefficient is given by the fraction of the coefficient of determination of the oil price forecast, the actual oil price ($R^2_{Forecast,Actual}$), the coefficient of determination of the oil price forecast, and the actual forecast at the date of forecast ($R^2_{Forecast,Actual(ForecastDate)}$). As the authors highlight, a TOTA coefficient < 1 can be seen as statistical evidence for trend adjustment behavior of the forecasters.

$$TOTA = \frac{R^2_{Forecast,\,Actual}}{R^2_{Forecast,\,Actual(ForecastDate)}}. \quad (6)$$

4.6 | Theil's U measure

Finally, we investigate the oil price forecast by applying the Theil's U measure (see Theil, 1955; as well as Andres & Spiwoks, 1999) as the fraction of the RMSE of the naïve forecast and the Brent forecasts $\left(U^{r}_{Brent}\right)$ and the naïve forecast and the WTI forecasts (U_{Brentr}), respectively.

$$U^{j}_{Brent} = \frac{RMSE^{j}_{Brent}}{RMSE^{Naive}_{Brent}} \text{ and } U^{k}_{WTI} = \frac{RMSE^{k}_{WTI}}{RMSE^{Naive}_{WTI}}. \quad (7)$$

A Theil's U measure (see Formula 7 above) of 1 signals that the forecast of institution *j* and *k*, respectively, is as good as the naïve prediction. A forecast with a Theil's U score below 1 performs better than the naïve prediction. In case of a Theil's U higher than 1, the naïve prediction outperforms the individual forecasts. Finally, it is important to note that the adjusted RMSE is not necessary in this context. When calculating the Theil's U score, only complete pairs of individual forecasts and the naïve prediction have been considered. Hence, a weighting of forecast errors is not necessary when calculating the Theil's U score.

5 | EVALUATION RESULTS

We commence this section by presenting and discussing the evaluation results from the sign accuracy test, the test for unbiasedness, and the test for forecast encompassing. Subsequently, we will present and briefly discuss the results of the forecast accuracy measures, both with and without controlling for the forecast horizon.

The results of the sign accuracy test, the test for unbiasedness, and the test for forecast encompassing are summarized in Tables 4 (Brent) and 5 (WTI) below. More detailed results can be found in Tables A1, A2, A3, A4, A5, A6 in the Appendix.

Regarding the sign accuracy test, we find empirical evidence for the individual forecasts being better than a random prediction (+) for

TABLE 4 Summary sign accuracy, unbiasedness, and forecast encompassing—Brent

Forecast	Sign accuracy	Unbiasedness	Encompassing
MEDIAN	NA	–	–
F1	+	–	–
F2	NA	–	–
F3	+	–	–
F4	+	–	+
F5	NA	–	+
F6	NA	–	+
F7	NA	–	+
F8	NA	–	+
F9	–	–	+
F10	NA	–	+
F11	NA	–	–
F12	+	–	+
F13	NA	–	+
F14	NA	–	–
F15	NA	–	+
F16	+	–	+
F17	+	–	+
F18	NA	–	–
F19	NA	–	+
F20	+	–	+

(Continues)

TABLE 4 (Continued)

Forecast	Sign accuracy	Unbiasedness	Encompassing
F21	+	−	−
F22	NA	−	+
F23	−	−	+
F24	NA	−	+
F25	+	−	+
F26	+	−	+
F27	NA	−	−
F28	+	−	+
F29	+	−	−
F30	+	−	−
F31	+	−	+
F32	NA	−	−
F33	NA	−	+
F34	NA	−	+
F35	NA	−	+
F36	+	−	−
F37	+	−	+
F38	+	−	−
F39	NA	−	−
F40	+	−	−
F41	+	−	+
F42	NA	−	−
F43	NA	−	−
F44	−	−	+
F45	NA	−	+
F46	+	−	−
F47	NA	−	−
F48	+	−	−
F49	NA	−	+
F50	NA	−	−
F51	NA	−	−
F52	+	−	−

Note. NA = no significant result; + = significant accurate prediction of sign; statistically significant unbiased, and statistically significant provision of additional information; − = no significant accurate prediction of sign, statistically significant biased, and statistically significant no provision of additional information.

TABLE 5 Summary sign accuracy, unbiasedness, and forecast encompassing—WTI

Forecast	Sign accuracy	Unbiasedness	Encompassing
MEDIAN	+	−	−
F1	+	−	+
F2	+	−	−
F3	+	−	−
F4	+	−	−
F5	+	−	−
F6	NA	−	−
F7	NA	−	+
F8	+	−	+
F9	NA	−	+
F10	−	−	−
F11	NA	−	−
F12	+	−	+
F13	+	−	+
F14	+	−	+
F15	NA	−	−
F16	NA	−	+
F17	+	−	+
F18	NA	−	+
F19	NA	+	−
F21	+	−	−
F22	+	−	+
F24	NA	−	+
F25	+	−	−
F26	+	−	+
F27	+	+	−
F28	−	−	−
F29	+	−	+
F30	+	−	−
F31	NA	−	−
F32	NA	−	−
F33	NA	−	−
F35	+	+	+
F37	+	−	+
F38	+	+	+
F39	+	−	−
F40	+	−	−
F41	NA	−	−
F43	NA	−	−
F44	−	−	−
F45	NA	−	+
F46	NA	−	+
F47	+	−	+
F48	+	−	−
F49	+	−	−
F50	NA	−	+
F53	NA	−	−
F54	NA	+	−
F51	NA	−	−
F52	+	−	−

Note. NA = no significant result; + = significant accurate prediction of sign, statistically significant unbiased, and statistically significant provision of additional information; − = no significant accurate prediction of sign, statistically significant biased, and statistically significant no provision of additional information.

22 (Brent) and 27 institutions (WTI). For both Brent and WTI, three institutions performed worse than a random prediction (−) for the direction of change. However, 28 (Brent) and 20 (WTI) delivered forecasts are not statistically significant different from a random prediction (NA), and hence also give no additional informational value, at least in terms of sign accuracy. The accuracy of the median forecast is not significantly different from a random prediction in the case of Brent.

This is an important result because the median forecast is often preferred by economic agents obtaining forecasts from external providers (e.g., portfolio managers). By contrast, the WTI median forecast passes the sign accuracy test.

Checking for unbiasedness of the forecasts leads to an even more disillusioning finding. For all predictions regarding the change of the price of Brent, the test for unbiasedness reveals that the forecasts are biased (−). In the case of the WTI predictions, following the regression approach of the test for unbiasedness for five forecasts, the conditions that α equals zero and β equals one seem to hold. However, for four out of these five forecasts, the widely acknowledged Durbin Watson test (see Durbin & Watson, 1950, 1951) delivers strong indications for the autocorrelation of the residuals. Hence, for these forecasts, on the other hand, the test for forecast encompassing delivers more encouraging results. We find empirical evidence for additional information (+) for 29 (Brent) and 21 (WTI) forecasts.

Summing up these heterogeneous intermediate results does not prompt the conclusion that forecasts delivered by professional analysts are unworkable. However, economic agents should also not unconditionally follow external forecasts. Given the results of the tests of forecast encompassing at least makes the case that one should not solely stick to the naïve prediction, because professional forecasters deliver additional information. Having said that, it does not allow any conclusions regarding forecasters' superiority over the naïve prediction. Using the forecast accuracy measures, we will be able to directly compare individual (but anonymized) forecasts with each other, the median forecast, and the naïve prediction. The results from the simple forecast accuracy measures are given in Tables 6 (Brent) and 7 (WTI) below.

Our findings regarding the measures of forecast accuracy are heterogeneous as well. For the predictions of both Brent and WTI, the means of the forecast errors have different signs and scales. The median forecast proves to be a rather accurate predictor regarding the mean forecast error. However, in the case of WTI, the naïve prediction obviously outperforms the median forecast and all individual forecasts. For the forecasts for both the price of Brent and WTI, the aRMSE as well as Theil's U measure reveal that a large fraction of the individual predictions are less accurate than the naïve prediction. These findings can also be seen as implications for topically oriented trend adjustments, which are empirically confirmed by the results of the TOTA coefficient. As stated above, a TOTA coefficient below 1 has to be seen as empirical evidence of trend adjustments. With the exception of one institution in the case of Brent and one institution for the WTI forecasts, we find empirical evidence of topically oriented trend adjustment for all individual forecasts and the median forecast.

As discussed earlier, thus far no effects regarding the forecast horizons have been considered. Because forecasts' accuracy should improve with data availability, forecast errors should be lower for shorter forecast horizons (see, e.g., Baghestani, 2010aa). We recall that without controlling for the forecast horizon in the case of Brent, the median forecast is not significantly different from a random prediction, is biased, and does not provide additional information. The median forecast for WTI outperforms a random forecast regarding sign accuracy, but is also biased and does not provide additional information. In Tables 8 and 9 below, the results of the three tests controlling for

TABLE 6 Results from simple accuracy measures (Brent)

Forecast	ME	aRMSE	TOTA	U	Obs
NAIV	1.35	16.09	NA	1.00	468
MEDIAN	0.27	11.72	0.25	1.00	468
F1	6.64	19.96	0.68	0.83	163
F2	4.35	11.57	0.41	0.90	316
F3	4.42	19.17	0.23	1.05	238
F4	−7.35	18.61	0.37	0.92	275
F5	−13.30	22.71	0.30	1.09	248
F6	6.21	17.46	0.27	1.09	361
F7	−15.14	16.17	0.58	0.83	60
F8	−8.18	17.10	0.00	1.16	156
F9	−2.43	19.28	0.07	1.24	186
F10	−0.57	20.92	0.22	1.09	415
F11	−2.55	20.58	0.40	0.91	270
F12	3.77	17.79	0.65	0.78	197
F13	5.02	14.93	0.49	0.83	307
F14	−2.25	15.65	0.01	1.21	215
F15	−5.07	22.11	0.19	1.11	269
F16	−9.46	13.08	0.00	1.19	132
F17	−3.57	13.66	0.32	0.97	249
F18	7.41	26.56	0.12	1.22	128
F19	−1.40	15.24	0.24	1.00	353
F20	−5.39	15.85	0.01	1.28	205
F21	−6.85	14.22	0.29	0.98	295
F22	−3.35	14.29	0.20	1.04	329
F23	0.88	14.02	0.22	1.06	363
F24	−4.16	17.54	0.02	1.24	262
F25	−4.67	26.58	0.28	1.05	337
F26	0.24	19.40	0.61	0.73	137
F27	0.60	12.96	0.26	0.99	415
F28	−3.78	19.30	0.27	0.96	327
F29	−4.90	17.35	0.16	1.02	371
F30	−2.74	10.54	0.00	1.33	70
F31	2.82	14.37	0.50	0.51	50
F32	−1.24	15.01	0.22	1.06	360
F33	−5.26	19.44	0.00	1.17	87
F34	11.32	18.01	0.40	1.15	52
F35	−5.09	14.27	0.00	1.36	127
F36	2.23	14.94	0.24	1.07	433
F37	0.31	15.54	0.29	1.00	421
F38	−2.83	16.20	0.29	0.97	383
F39	−10.55	15.88	0.00	1.24	143
F40	0.92	15.62	0.28	1.01	419
F41	3.38	14.79	4.12	0.44	68
F42	5.45	12.07	1.10	0.64	43
F43	2.12	16.24	0.25	1.07	356
F44	13.04	14.45	0.64	0.93	70
F45	10.91	17.64	0.21	1.25	259
F46	−0.97	18.29	0.13	1.13	345
F47	3.38	15.71	0.35	0.97	325
F48	−4.75	19.44	0.00	0.55	110
F49	−9.64	13.86	0.12	1.08	240

(Continues)

TABLE 6 (Continued)

Forecast	ME	aRMSE	TOTA	U	Obs
F50	21.89	16.01	0.01	1.99	59
F51	−21.76	11.94	0.02	1.41	45
F52	−10.68	16.46	0.03	0.90	44

Note. ME = mean forecast error; aRMSE = Adjusted root mean squared error.

TABLE 7 Results from simple accuracy measures (WTI)

Forecast	ME	aRMSE	TOTA	U	Obs
NAIV	−0.34	16.98	NA	1.00	468
MEDIAN	0.96	9.88	0.16	0.94	468
F1	6.77	23.63	0.64	0.76	163
F2	3.04	25.55	0.28	0.86	324
F3	6.18	105.12	0.13	1.00	234
F4	−9.21	15.19	0.25	0.84	173
F5	−15.83	10.75	0.06	1.20	160
F6	6.31	17.86	0.18	1.07	365
F7	−12.06	12.36	2.31	0.70	60
F8	−9.00	18.86	0.01	1.18	156
F9	−3.28	14.61	0.04	1.22	178
F10	3.78	15.47	0.14	1.06	355
F11	−1.38	13.56	0.37	0.80	294
F12	4.10	15.24	0.63	0.73	193
F13	9.08	18.56	0.49	0.83	265
F14	−0.89	17.70	0.00	1.14	215
F15	−4.25	15.69	0.08	1.05	276
F16	−9.39	13.61	0.00	1.18	132
F17	−2.64	12.38	0.21	0.91	242
F18	7.63	19.48	0.05	1.22	133
F19	0.23	11.29	0.00	0.54	44
F21	−7.58	11.30	0.03	1.05	235
F22	−2.58	25.83	0.10	0.96	321
F24	−5.25	14.03	0.01	1.20	246
F25	−2.97	12.30	0.30	0.90	349
F26	4.00	11.71	0.79	0.63	137
F27	−0.57	15.09	0.00	1.28	122
F28	−6.99	17.26	0.11	0.94	159
F29	−2.13	12.98	0.04	0.97	370
F30	−4.97	10.58	0.00	1.33	70
F31	5.74	15.41	0.23	0.52	54
F32	−0.96	11.97	0.11	1.00	363
F33	−5.31	8.99	0.00	1.18	88
F35	−3.52	13.03	0.00	1.33	120
F37	5.24	13.99	0.32	0.85	365
F38	−2.07	16.27	0.12	0.92	363
F39	−11.38	12.15	0.00	1.23	143
F40	1.97	12.65	0.17	0.97	423
F41	1.06	12.09	0.00	0.42	68
F43	2.49	13.34	0.19	1.02	356
F44	10.01	12.28	0.57	0.86	68
F45	10.57	11.62	0.11	1.19	260
F46	−0.50	14.35	0.05	1.06	334
F47	2.74	11.43	0.22	0.92	317

(Continues)

TABLE 7 (Continued)

Forecast	ME	aRMSE	TOTA	U	Obs
F48	2.02	16.11	0.52	0.45	110
F49	−7.36	15.76	0.03	1.02	220
F50	36.43	13.15	0.01	2.27	42
F53	3.38	11.69	0.34	0.46	29
F54	−2.43	11.99	0.07	0.46	36
F51	−23.27	14.89	0.02	1.45	45
F52	−9.47	18.68	0.00	0.68	44

Note. WTI = Western Texas Intermediate; ME = mean forecast error; aRMSE = Adjusted root mean squared error.

TABLE 8 Summary sign accuracy, unbiasedness, and forecast encompassing controlling for the forecast horizon—Brent

Forecast horizon	Sign accuracy	Unbiasedness	Encompassing
ALL	**NA**	**−**	**−**
H1	NA	−	+
H2	−	−	+
H3	NA	+	+
H4	NA	+	+
H5	NA	+	−
H6	NA	+	−
H7	NA	+	−
H8	−	−	−
H9	NA	+	−
H10	NA	+	−
H11	NA	+	−
H12	NA	+	−

Note. NA = no significant result; + = significant accurate prediction of sign, statistically significant unbiased, and statistically significant provision of additional information; − = no significant accurate prediction of sign, statistically significant biased, and statistically significant no provision of additional information; ALL = not controlling for the forecast horizon.

TABLE 9 Summary sign accuracy, unbiasedness, and forecast encompassing controlling for the forecast horizon—WTI

Forecast horizon	Sign accuracy	Unbiasedness	Encompassing
ALL	**+**	**−**	**−**
H1	NA	−	−
H2	NA	−	−
H3	NA	+	+
H4	NA	+	−
H5	−	+	−
H6	−	+	−
H7	NA	+	−
H8	NA	+	−
H9	NA	+	−
H10	−	+	−
H11	−	+	−
H12	NA	+	−

Note. NA = no significant result; + = significant accurate prediction of sign, statistically significant unbiased, and statistically significant provision of additional information; − = no significant accurate prediction of sign, statistically significant biased, and statistically significant no provision of additional information; ALL = not controlling for the forecast horizon; WTI = Western Texas Intermediate.

the forecast horizon are shown. Following these results, for the median forecast for the price of both Brent and WTI, the random prediction either outperformed the median forecast in terms of sign accuracy or no statistical evidence of statistically significant differences in sign accuracy were found. In the case of the WTI forecast, this result departs from the earlier findings when not controlling for the forecast horizon. Regarding the test for unbiasedness, differentiating between the 12 different forecast horizons shows that for longer-term predictions the median forecast is generally unbiased for both Brent and WTI.[2] The corresponding results for the test for forecast encompassing are mixed. In the case of the Brent price forecast, it can be stated that at least for shorter forecast horizons, the median prediction provides additional information. In terms of the test for forecast encompassing, the results only slightly deviate. Most noteworthy in this context is the finding that in the case of Brent, the median forecast seems to provide additional information, especially for shorter horizons.

Controlling for the forecast horizon for the simple forecast accuracy measures also delivers noteworthy empirical results. The results are shown in Figures 5 and 6. We use box plots, which are quite common in the context of forecast evaluation (see, e.g., Bowles et al., 2007; Bowles et al., 2010; Fildes, 1992; as well as Davydenko & Fildes, 2013 for applications in the context of survey predictions), because they deliver a useful depiction of large numbers of individual forecasts in total and for each forecast horizon. As noted above, uncertainty is a critical aspect for decision-makers. The straightforward interpretation of the box plots delivers academic researchers and decision makers a useful first impression of the underlying data and forecasts, respectively. Given the large amount of individual forecasts, this is a clear advantage.

The box plot figures (without outliers) for the mean forecast error and the aRMSE for Brent and WTI (Figure 5) show that the mean forecast errors only slightly increase with the forecast horizon. Interestingly, the dispersion of forecasts and mean forecast errors increases with a rising forecast horizon, which indicates rising uncertainty. It is noteworthy that the mean of the forecast errors seems be centered on zero for the given forecast horizons. Hence, there does not seem to be a majority of forecasters over- or under-estimating the WTI or Brent price on average. Given the large oil price swings during the period under investigation, this result is somewhat surprising. Furthermore, the aRMSE measures also increase with the forecast horizon. Despite being foreseeable, the implication of this result is important when assessing the usefulness of oil price predictions of professional forecasters.

The results of the forecast horizon-dependent TOTA coefficient and Theil's U measure for Brent and WTI are shown in Figure 6 (box plot without outliers). In the case of the TOTA coefficient, we see rising topically oriented trend adjustment with rising forecast horizons. In turn, this means that for longer forecast horizons, analysts seem to be more status quo-dependent. Notwithstanding, only a small fraction of forecasters in the case of the price of WTI and even fewer in the case of Brent fulfill the requirement of no topically oriented trend adjustment, even when considering short forecast horizons. In terms of the Theil's U measure, the results are somewhat different. Especially for the short-term forecasts (i.e., Horizon 1), the high Theil's U measure reveals the superiority of the naïve prediction against the majority of the individual forecasts. At least with a rising forecast horizon, a growing fraction of individual forecasters outperforms the simple naïve prediction. However, it is important to note that in the case of the median Brent forecast, the results from the horizon-dependent tests for forecast encompassing indicate that relevant additional information is delivered in the short-term. Overall, the Figures 5 and 6 demonstrate quite clearly the limitations of survey forecasts. This does not necessarily mean that survey forecasts for the price of oil deliver no value added. However, knowing the specific constraints of the professional forecasts is inevitably necessary when using oil price forecasts as a methodology for decision-support.

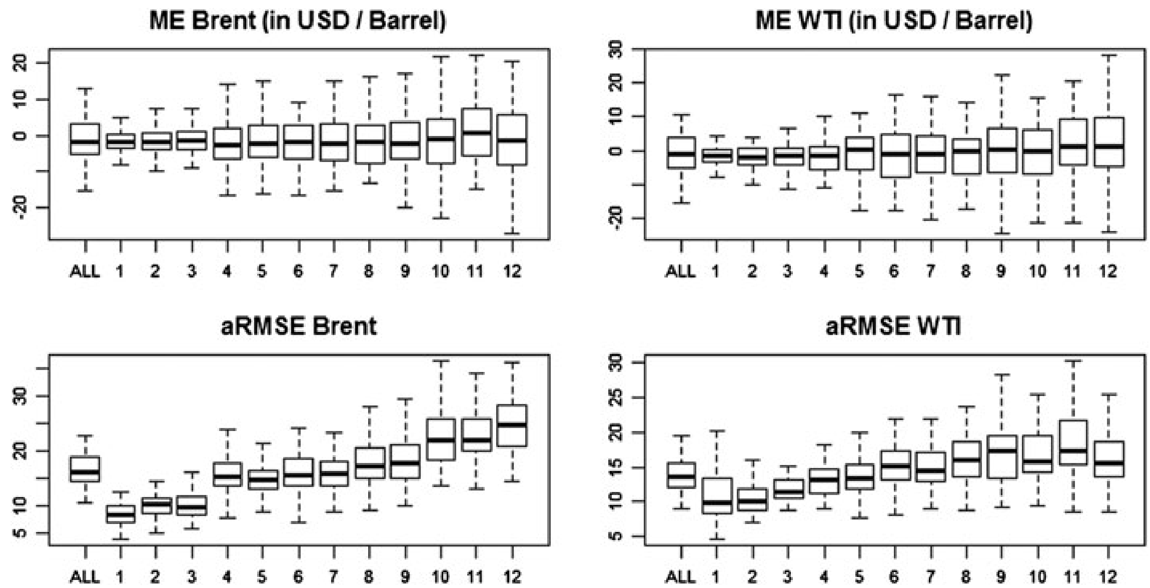


FIGURE 5 Box plots for errors measures mean forecast error (ME) and adjusted root mean squared forecast error (aRMSE) for Brent and Western Texas Intermediate (WTI) regarding the forecast horizon 1 to 12 months; ALL = not controlling for the forecast horizon

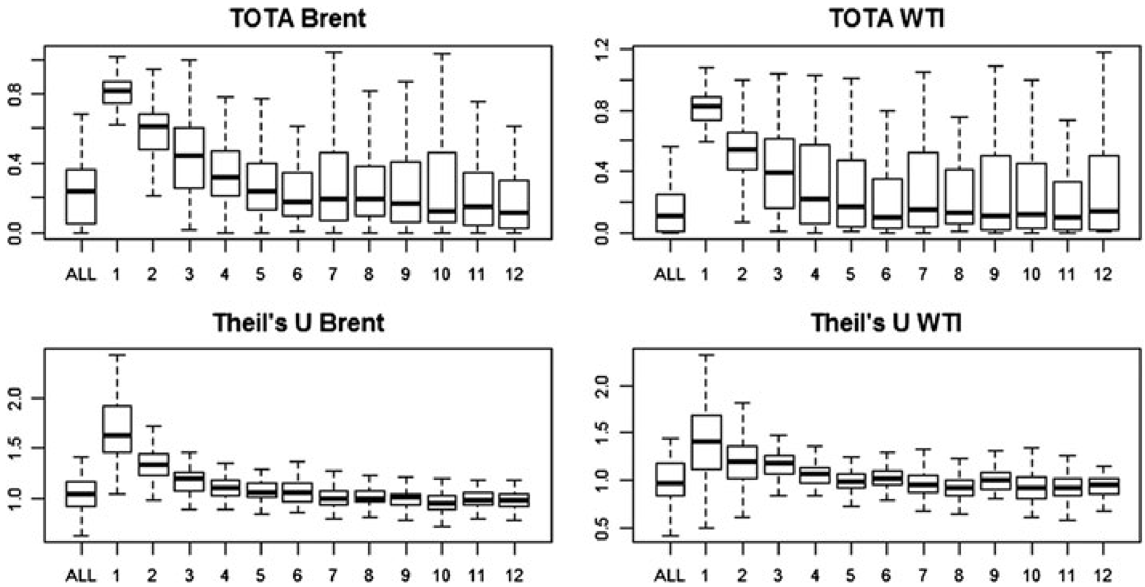


FIGURE 6 Box plots for TOTA coefficient and Theil's U measure for Brent and Western Texas Intermediate (WTI) regarding the forecast horizon 1 to 12 months; ALL = not controlling for the forecast horizon

6 | SUMMARY AND CONCLUSIONS

We have empirically investigated survey forecasts for the globally most relevant crude oil benchmarks. Using a novel comprehensive disaggregated dataset incorporating individual forecasts for both Brent and WTI, we have applied a standard framework for forecast evaluation with and without controlling for the forecast horizon. We performed tests for unbiasedness, sign accuracy, and forecast encompassing. As accuracy measures, mean forecast errors, an aRMSE, the TOTA coefficient, and the Theil's U measure have been used.

We have found mixed evidence regarding the rationality, additional information and accuracy of the crude oil price forecasts. In the case of tests for unbiasedness, sign accuracy, and forecast encompassing, we even find contracting results, at least regarding the median forecast. Referring to the forecast accuracy measures, one major finding is the strong relationship between the oil price at the date of forecast delivery and the prediction of the professional analysts. A large fraction of the forecasters tend to follow the naïve prediction. In fact, the TOTA coefficient shows that the topically oriented trend adjustments becomes more pronounced with a rising forecast horizon for the price of both Brent and WTI. Interestingly, a growing fraction of analysts perform better than the naïve prediction for longer-term forecast horizons, which in general contributes to the usefulness of the forecasts.

However, given the status quo dependence and bearing in mind the empirical literature dealing with other financial market forecasts (e.g., interest rates), it can be stated that crude oil predictions share common characteristics, and hence do not stand out against these other financial market variables.

Financial market forecasts in general and especially oil price forecasts are important decision-support methodologies. Hence, for decision-makers, these results are noteworthy in several respects. Most importantly, when making use of survey forecasts, they have to bear in mind the dependency of forecast accuracy regarding the forecast horizon and the related consequences for the input costs and their value chain. Furthermore, the TOTA coefficients unveil a strong status quo bias. These may reduce the usefulness of professional forecasts in the context of strategic decision-making as a forward looking view is needed.

Further research should focus on decision-makers' implementation strategies of survey forecasts for crude oil prices that incorporate the accuracy of survey predictions. In this context, especially the relationship between uncertainty and forecasters' disagreement is an important research field.

ENDNOTES

[1] Because it is not the focus of this paper to analyze the forecast accuracy of specific individual forecasters, an anonymized data set will be presented throughout the paper.

[2] The results of the Durbin Watson test for autocorrelation (see, Durbin & Watson, 1950, 1951) deliver no evidence for autocorrelation of the residuals for the relevant forecasts. The detailed results are not reported here in order to preserve space.

ORCID

Frederik Kunze http://orcid.org/0000-0002-5843-6910

REFERENCES

Abosedra, S., & Baghestani, H. (2004). On the predictive accuracy of crude oil futures prices. *Energy Policy*, *32*(12), 1389–1393. https://doi.org/10.1016/S0301-4215(03)00104-6

Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). Forecasting the price of oil. *Handbook of Economic Forecasting*, *2*, 427–507.

Andres, P., & Spiwoks, M. (1999). Forecast quality matrix: A methodological survey of judging forecast quality of capital market forecasts. *Journal of Economics and Statistics*, *219*, 513–542.

Atalla, T., Joutz, F., & Pierru, A. (2016). Does disagreement among oil price forecasters reflect volatility? Evidence from the ECB surveys. *International Journal of Forecasting*, *32*(4), 1178–1192. https://doi.org/10.1016/j.ijforecast.2015.09.009

Audretsch, D. B., & Stadtmann, G. (2005). Biases in FX-forecasts: Evidence from panel data. *Global Finance Journal*, *16*(1), 99–111. https://doi.org/10.1016/j.gfj.2005.05.001

Baghestani, H. (2008). Federal reserve versus private information: Who is the best unemployment rate predictor? *Journal of Policy Modeling*, *30*(1), 101–110. https://doi.org/10.1016/j.jpolmod.2007.05.002

Baghestani, H. (2010a). Forecasting the 10-year US treasury rate. *Journal of Forecasting*, *29*(8), 673–688. https://doi.org/10.1002/for.1165

Baghestani, H. (2010b). Evaluating blue chip forecasts of the trade-weighted dollar exchange rate. *Applied Financial Economics*, *20*(24), 1879–1889. https://doi.org/10.1080/09603107.2010.526578

Baghestani, H., Arzaghi, M., & Kaya, I. (2015). On the accuracy of blue chip forecasts of interest rates and country risk premiums. *Applied Economics*, *47*(2), 113–122. https://doi.org/10.1080/00036846.2014.959656

Banker, R. D., & Kauffman, R. J. (2004). 50th anniversary article: The evolution of research on information systems: A fiftieth-year survey of the literature in management science. *Management Science*, *50*(3), 281–298. https://doi.org/10.1287/mnsc.1040.0206

Barron, M., & Targett, D. (1988). Managing forecasts: User and expert. *Management Decision*, *26*(1), 41–44. https://doi.org/10.1108/eb001485

Basher, S. A., & Sadorsky, P. (2006). Oil price risk and emerging stock markets. *Global Finance Journal*, *17*(2), 224–251. https://doi.org/10.1016/j.gfj.2006.04.001

Beckmann, J., & Czudaj, R. (2017). The impact of uncertainty on professional exchange rate forecasts. *Journal of International Money and Finance*, *73*, 296–316. https://doi.org/10.1016/j.jimonfin.2017.02.009

Beechey, M., & Österholm, P. (2014). Policy interest-rate expectations in Sweden: A forecast evaluation. *Applied Economics Letters*, *21*(14), 984–991. https://doi.org/10.1080/13504851.2014.904480

Behmiri, N. B., & Pires Manso, J. R. (2013). Crude oil price forecasting techniques: A comprehensive review of literature. *CAIA Alternative Investment Analyst Review*, *2*(3), 30–48. https://doi.org/10.2139/ssrn.2275428

Belongia, M. T. (1987). Predicting interest rates: A comparison of professional and market-based forecasts. Federal Reserve Bank of St. Louis review, March 1987.

Benke, H. (2006). Was leisten Kapitalmarktprognosen? – Die Sicht eines Stiftungsmanagers. *Kreditwesen*, *17*, 902–906.

Bentzen, J. (2007). Does OPEC influence crude oil prices? Testing for co-movements and causality between regional crude oil prices. *Applied Economics*, *39*(11), 1375–1385. https://doi.org/10.1080/00036840600606344

Bofinger, P., & Schmidt, R. (2003). On the reliability of professional exchange rate forecasts: An empirical analysis for the€/US-$ rate. *Financial Markets and Portfolio Management*, *17*(4), 437–449. https://doi.org/10.1007/s11408-003-0403-z

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., & Rautanen, T. (2007). The ECB survey of professional forecasters (SPF)—A review after eight years' experience. ECB Occasional Paper, No. 59. http://hdl.handle.net/10419/154512

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., & Rautanen, T. (2010). An evaluation of the growth and unemployment forecasts in the ECB survey of professional forecasters. *OECD Journal. Journal of Business Cycle Measurement and Analysis*, *2010*(2), 63.

Chen, Q., Costantini, M., & Deschamps, B. (2016). How accurate are professional forecasts in Asia? Evidence from ten countries. *International Journal of Forecasting*, *32*(1), 154–167. https://doi.org/10.1016/j.ijforecast.2015.05.004

Clements, M. P., & Harvey, D. I. (2010). Forecast encompassing tests and probability forecasts. *Journal of Applied Econometrics*, *25*(6), 1028–1062. https://doi.org/10.1002/jae.1097

Clements, M. P. (2014). Forecast uncertainty—Ex ante and ex post: US inflation and output growth. *Journal of Business & Economic Statistics*, *32*(2), 206–216. https://doi.org/10.1080/07350015.2013.859618

Davydenko, A., & Fildes, R. (2013). Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. *International Journal of Forecasting*, *29*(3), 510–522. https://doi.org/10.1016/j.ijforecast.2012.09.002

Diebold, F. X., & Lopez, J. A. (1996). *Forecast evaluation and combination*. (pp. 1–46). University of Pennsylvania: mimeo, Department of Economics.

Dominguez, K. M. (1986). Are foreign exchange forecasts rational?: New evidence from survey data. *Economics Letters*, *21*(3), 277–281. https://doi.org/10.1016/0165-1765(86)90188-6

Durbin, J., & Watson, G. (1950). Testing for serial correlation in least squares regression. I. *Biometrika*, *37*, 409–428. https://doi.org/10.1093/biomet/37.3-4.409

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, *38*, 159–177. https://doi.org/10.1093/biomet/38.1-2.159

Espinoza, R., Fornari, F., & Lombardi, M. J. (2012). The role of financial variables in predicting economic activity. *Journal of Forecasting*, *31*(1), 15–46. https://doi.org/10.1002/for.1212

Fildes, R. (1992). The evaluation of extrapolative forecasting methods. *International Journal of Forecasting*, *8*(1), 81–98. https://doi.org/10.1016/0169-2070(92)90009-X

Frey, G., Manera, M., Markandya, A., & Scarpa, E. (2009). Econometric models for oil price forecasting: A critical survey. *CESifo Forum*, *1*, 29–44. Ifo Institute for Economic Research at the University of Munich

Friedman, B. M. (1980). Survey evidence on the "rationality" of interest rate expectations. *Journal of Monetary Economics*, *6*(4), 453–465. https://doi.org/10.1016/0304-3932(80)90001-X

Glas, A., & Hartmann, M. (2016). Inflation uncertainty, disagreement and monetary policy: Evidence from the ECB survey of professional forecasters. *Journal of Empirical Finance*, *39*, 215–228. https://doi.org/10.1016/j.jempfin.2016.05.001

Greer, M. (2003). Directional accuracy tests of long-term interest rate forecasts. *International Journal of Forecasting*, *19*(2), 291–298. https://doi.org/10.1016/S0169-2070(01)00141-8

Gripaios, P. (1994). The use and abuse of economic forecasts. *Management Decision*, *32*(6), 61–64. https://doi.org/10.1108/00251749410065178

Haushalter, G. D., Heron, R. A., & Lie, E. (2002). Price uncertainty and corporate value. *Journal of Corporate Finance*, *8*(3), 271–286. https://doi.org/10.1016/S0929-1199(01)00043-8

He, Y., Shouyang, W., & Lai, K. K. (2010). Global economic activity and crude oil prices: A cointegration analysis. *Energy Economics*, *32*(4), 868–876. https://doi.org/10.1016/j.eneco.2009.12.005

Herwartz, H., & Schlüter, S. (2017). On the predictive information of futures' prices: A wavelet-based assessment. *Journal of Forecasting*, *36*(4), 345–356. https://doi.org/10.1002/for.2435

Holden, K., & Peel, D. A. (1990). On testing for unbiasedness and efficiency of forecasts. *The Manchester School*, *58*(2), 120–127. https://doi.org/10.1111/j.1467-9957.1990.tb00413.x

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Ince, O., & Molodtsova, T. (2017). Rationality and forecasting accuracy of exchange rate expectations: Evidence from survey-based forecasts. *Journal of International Financial Markets, Institutions & Money*, *47*, 131–151. https://doi.org/10.1016/j.intfin.2016.11.002

Jongen, R., Verschoor, W. F., Wolff, C. C., & Zwinkels, R. C. (2012). Explaining dispersion in foreign exchange expectations: A heterogeneous agent approach. *Journal of Economic Dynamics and Control*, *36*(5), 719–735. https://doi.org/10.1016/j.jedc.2012.01.005

Kaufmann, R. K., Dees, S., Karadeloglou, P., & Sanchez, M. (2004). Does OPEC matter? An econometric analysis of oil prices. *The Energy Journal*, *25*(4), 67–90. https://doi.org/10.2307/41323358

Kaufmann, R. K., & Ullman, B. (2009). Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices.

Energy Economics, *31*(4), 550–558. https://doi.org/10.1016/j.eneco.2009.01.013

Kenny, G., Kostka, T., & Masera, F. (2015). Can macroeconomists forecast risk? Event-based evidence from the euro-area SPF. *International Journal of Central Banking*, *11*(4), 1–46.

Kilian, L. (2008). Exogenous oil supply shocks: How big are they and how much do they matter for the US economy? *The Review of Economics and Statistics*, *90*(2), 216–240. Doi:10.1162/rest.90.2.216 doi:https://doi.org/10.1162/rest.90.2.216

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, *99*, 1053–1069.

Kisinbay, T. (2010). The use of encompassing tests for forecast combinations. *Journal of Forecasting*, *29*(8), 715–727. https://doi.org/10.1002/for.1170

Knetsch, T. A. (2007). Forecasting the price of crude oil via convenience yield predictions. *Journal of Forecasting*, *26*(7), 527–549. https://doi.org/10.1002/for.1040

Kolb, R. A., & Stekler, H. O. (1996). How well do analysts forecast interest rates? *Journal of Forecasting*, *15*, 385–394. https://doi.org/10.1002/(SICI)1099-131X(199609)15:5<385::AID-FOR627>3.0.CO;2-B

Kunze, F. (2014). The European financial crisis—a challenge for ten-year German government bond yield forecasts? *International Journal of Bonds and Derivatives*, *1*(2), 171–185. https://doi.org/10.1504/IJBD.2014.067402

Kunze, F., Gruppe, M., & Wendler, T. (2015). Interest rate forecasting and the financial crisis: a turning point in more than just one way. *International Journal of Financial Engineering and Risk Management*, *2*(1), 1–16. https://doi.org/10.1504/IJFERM.2015.068850

Lahiri, K., & Sheng, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics*, *25*(4), 514–538. https://doi.org/10.1002/jae.1167

Lawrence, D. B. (1991). Managerial evaluation of exogenous forecast sources. *Managerial and Decision Economics*, *12*(3), 249–259. https://doi.org/10.1002/mde.4090120307

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: Profits versus the conventional error measures. *The American Economic Review*, *81*(3), 580–590.

Lowe, E. A. (1970). Efficient forecasting and management information systems. *Management Decision*, *4*(4), 51–54. https://doi.org/10.1108/eb000950

MacDonald, M., & Marsh, I. (1993). On the efficiency of oil price forecasts. *Applied Financial Economics*, *3*(4), 293–302.

MacDonald, R. (1990). Are foreign exchange market forecasters “rational”: Some survey-based tests. *The Manchester School*, *58*(3), 229–241. https://doi.org/10.1111/j.1467-9957.1990.tb00420.x

McCracken, M. W., & West, K. D. (2008). Inference about predictive ability. In M. P. Clements, & D. F. Hendry (Eds.), *A companion to economic forecasting* (pp. 299–321). Oxford: Blackwells.

Meub, L., Proeger, T., Bizer, K., & Spiwoks, M. (2015). Strategic coordination in forecasting–an experimental study. *Finance Research Letters*, *13*, 155–162.

Miller, J. I., & Ratti, R. A. (2009). Crude oil and stock markets: Stability, instability, and bubbles. *Energy Economics*, *31*(4), 559–568. https://doi.org/10.1016/j.eneco.2009.01.009

Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic forecasts and expectation* (pp. 3–46). New York: Columbia University Press.

Mitchell, K., & Pearce, D. K. (2007). Professional forecasts of interest rates and exchange rates: Evidence from the wall street journal's panel of economists. *Journal of Macroeconomics*, *29*(4), 840–854. https://doi.org/10.1016/j.jmacro.2005.11.004

Natal, J. (2012). Monetary policy response to oil price shocks. *Journal of Money, Credit and Banking*, *44*(1), 53–101.

Pesaran, M. H., & Weale, M. (2006). Survey expectations. In *Handbook of economic forecasting* (Vol. 1) (pp. 715–776). https://doi.org/10.1016/S1574-0706(05)01014-1

Pierdzioch, C., Rülke, C. J., & Stadtmann, G. (2010). New evidence of anti-herding of oil-price forecasters. *Energy Economics*, *32*(6), 1456–1459.

Pierdzioch, C., Rülke, C. J., & Stadtmann, G. (2013). Oil price forecasting under asymmetric loss. *Applied Economics*, *45*(17), 2371–2379.

Pierdzioch, C., & Rülke, J. C. (2015). On the directional accuracy of forecasts of emerging market exchange rates. *International Review of Economics & Finance*, *38*, 369–376. https://doi.org/10.1016/j.iref.2015.03.003

Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *The Quarterly Review of Economics and Finance*, *51*(3), 236–247.

Reitz, S., Rülke, J. C., & Stadtmann, G. (2009). *Are oil price forecasters finally right? Regressive expectations toward more fundamental values of the oil price*, Deutsche Bundesbank Discussion Papers, *Series 1: Economic Studies*, 32, 2009.

Reitz, S., Stadtmann, G., & Taylor, M. P. (2010). The effects of Japanese interventions on FX-forecast heterogeneity. *Economics Letters*, *108*(1), 62–64. https://doi.org/10.1016/j.econlet.2010.04.007

Rötheli, T. F. (1998). Forecasting among alternative strategies in the management of uncertainty. *Managerial and Decision Economics*, *19*(3), 179–187. https://doi.org/10.1002/(SICI)1099-1468(199805)19:3<179::AID-MDE878>3.0.CO;2-K

Rülke, J. C., Pierdzioch, C., & Stadtmann, G. (2012). On the internal consistency of short-term, medium-term and long-term oil price forecasts. *Applied Economics*, *44*(21), 2757–2765.

Ryan, L., & Whiting, B. (2017). Multi-model forecasts of the west Texas intermediate crude oil spot price. *Journal of Forecasting*, *36*(4), 395–406. https://doi.org/10.1002/for.2440

Sadorsky, P. (2008). Assessing the impact of oil prices on firms of different sizes: Its tough being in the middle. *Energy Policy*, *36*(10), 3854–3861. https://doi.org/10.1016/j.enpol.2008.07.019

Schmidbauer, H., & Rösch, A. (2012). OPEC news announcements: Effects on oil price expectation and volatility. *Energy Economics*, *34*(5), 1656–1663. https://doi.org/10.1016/j.eneco.2012.01.006

Singleton, K. J. (2013). Investor flows and the 2008 boom/bust in oil prices. *Management Science*, *60*(2), 300–318. https://doi.org/10.1287/mnsc.2013.1756

Spiwoks, M., Bedke, N., & Hein, O. (2008). Forecasting the past: The case of US interest rate forecasts. *Financial Markets and Portfolio Management*, *22*(4), 357–379. https://doi.org/10.1007/s11408-008-0087-5

Spiwoks, M., Bedke, N., & Hein, O. (2009). The pessimism of Swiss bond market analysts and the limits of the sign accuracy test-an empirical investigation of their forecasting success between 1998 and 2007. *International Bulletin of Business Administration*, *4*, 6–19.

Spiwoks, M., Bedke, N., & Hein, O. (2010). Topically orientated trend adjustment and autocorrelation of the residuals - an empirical investigation of the forecasting behavior of bond market analysts in Germany. *Journal of Money, Investment, and Banking*, *14*, 16–35.

Spiwoks, M., Scheier, J., & Hein, O. (2014). On assessing economic forecasts: An evaluation of predictions on GDP, Industrial Production and Private Consumer Spending Trends in Twelve Industrial Nations. *Advances in Management and Applied Economics*, *4*(2), 49.

Swanson, E. T. (2006). Have increases in federal reserve transparency improved private sector interest rate forecasts? *Journal of Money, Credit, and Banking*, *38*(3), 791–819.

Takagi, S. (1991). Exchange rate expectations: A survey of survey studies. *IMF Staff Papers*, *38*(1), 156–183. https://doi.org/10.2307/3867039

Theil, H. (1955). Who forecasts best?, International Economic Papers. 5, 194–199.

Waddell, D., & Sohal, A. S. (1994). Forecasting: The key to managerial decision making. *Management Decision*, *32*(1), 41–49. https://doi.org/10.1108/00251749410050697

Wegener, C., Basse, T., Kunze, F., & von Mettenheim, H. J. (2016). Oil prices and sovereign credit risk of oil producing countries: An empirical investigation. *Quantitative Finance*, *16*(12), 1961–1968. https://doi.org/10.1080/14697688.2016.1211801

Wegener, C., Spreckelsen, C., Basse, T., & Mettenheim, H. J. (2016). Forecasting government bond yields with neural networks considering cointegration. *Journal of Forecasting*, *35*(1), 86–92. https://doi.org/10.1002/for.2385

How to cite this article: Kunze F, Spiwoks M, Bizer K, Windels T. The usefulness of oil price forecasts—Evidence from survey predictions. *Manage Decis Econ*. 2018;1–20. https://doi.org/10.1002/mde.2916

APPENDIX

TABLE A1 Results sign accuracy test Brent

Forecast	Test statistic	Critical value 99%	Critical value 95%	Critical value 90%	Obs	Significance level	Result
MEDIAN	7.37	6.63	3.84	2.71	468	***	+
F1	0.02	6.63	3.84	2.71	163		NA
F2	3.65	6.63	3.84	2.71	316	*	+
F3	0.99	6.63	3.84	2.71	238		NA
F4	6.77	6.63	3.84	2.71	275	***	+
F5	3.01	6.63	3.84	2.71	248	*	+
F6	1.40	6.63	3.84	2.71	361		NA
F7	0.07	6.63	3.84	2.71	60		NA
F8	2.38	6.63	3.84	2.71	156		NA
F9	0.09	6.63	3.84	2.71	186		NA
F10	10.31	6.63	3.84	2.71	415	***	–
F11	0.26	6.63	3.84	2.71	270		NA
F12	0.29	6.63	3.84	2.71	197		NA
F13	15.63	6.63	3.84	2.71	307	***	+
F14	0.49	6.63	3.84	2.71	215		NA
F15	0.53	6.63	3.84	2.71	269		NA
F16	0.85	6.63	3.84	2.71	132		NA
F17	11.98	6.63	3.84	2.71	249	***	+
F18	3.51	6.63	3.84	2.71	128	*	+
F19	0.58	6.63	3.84	2.71	353		NA
F20	0.17	6.63	3.84	2.71	205		NA
F21	24.03	6.63	3.84	2.71	295	***	+
F22	3.13	6.63	3.84	2.71	329	*	+
F23	0.05	6.63	3.84	2.71	363		NA
F24	5.66	6.63	3.84	2.71	262	**	–
F25	0.01	6.63	3.84	2.71	337		NA
F26	3.69	6.63	3.84	2.71	137	*	+
F27	9.56	6.63	3.84	2.71	415	***	+
F28	0.01	6.63	3.84	2.71	327		NA
F29	27.01	6.63	3.84	2.71	371	***	+
F30	6.58	6.63	3.84	2.71	70	**	+
F31	2.90	6.63	3.84	2.71	50	*	+
F32	4.25	6.63	3.84	2.71	360	**	+
F33	1.51	6.63	3.84	2.71	87		NA
F34	1.84	6.63	3.84	2.71	52		NA
F35	1.02	6.63	3.84	2.71	127		NA
F36	2.54	6.63	3.84	2.71	433		NA
F37	5.49	6.63	3.84	2.71	421	**	+
F38	3.53	6.63	3.84	2.71	383	*	+
F39	4.37	6.63	3.84	2.71	143	**	+

(Continues)

TABLE A1 (Continued)

Forecast	Test statistic	Critical value 99%	Critical value 95%	Critical value 90%	Obs	Significance level	Result
F40	0.09	6.63	3.84	2.71	419		NA
F41	16.12	6.63	3.84	2.71	68	***	+
F42	3.95	6.63	3.84	2.71	43	**	+
F43	0.00	6.63	3.84	2.71	356		NA
F44	-	6.63	3.84	2.71	70		NA
F45	3.40	6.63	3.84	2.71	259	*	−
F46	0.12	6.63	3.84	2.71	345		NA
F47	4.77	6.63	3.84	2.71	325	**	+
F48	2.68	6.63	3.84	2.71	110		NA
F49	14.42	6.63	3.84	2.71	240	***	+
F50	0.41	6.63	3.84	2.71	59		NA
F51	1.98	6.63	3.84	2.71	45		NA
F52	0.48	6.63	3.84	2.71	44		NA

Note. + = better than a random forecast; − = worse than random forecast; NA = no statistically significant difference from a random forecast.
Significance levels:
*90%, **95%, ***99%.

TABLE A2 Results sign accuracy test WTI

Forecast	Test statistic	Critical value 99%	Critical value 95%	Critical value 90%	Obs	Significance level	Result
MEDIAN	9.15	6.63	3.84	2.71	468	***	+
F1	3.16	6.63	3.84	2.71	163	*	+
F2	6.53	6.63	3.84	2.71	324	**	+
F3	4.10	6.63	3.84	2.71	234	**	+
F4	11.91	6.63	3.84	2.71	173	***	+
F5	2.77	6.63	3.84	2.71	160	*	+
F6	1.79	6.63	3.84	2.71	365		NA
F7	0.01	6.63	3.84	2.71	60		NA
F8	7.91	6.63	3.84	2.71	156	***	+
F9	1.06	6.63	3.84	2.71	178		NA
F10	2.91	6.63	3.84	2.71	355	*	−
F11	2.61	6.63	3.84	2.71	294		NA
F12	8.41	6.63	3.84	2.71	193	***	+
F13	12.74	6.63	3.84	2.71	265	***	+
F14	3.78	6.63	3.84	2.71	215	*	+
F15	0.80	6.63	3.84	2.71	276		NA
F16	0.59	6.63	3.84	2.71	132		NA
F17	10.59	6.63	3.84	2.71	242	***	+
F18	1.75	6.63	3.84	2.71	133		NA
F19	0.00	6.63	3.84	2.71	44		NA
F21	23.24	6.63	3.84	2.71	235	***	+
F22	14.78	6.63	3.84	2.71	321	***	+
F24	0.00	6.63	3.84	2.71	246		NA
F25	4.57	6.63	3.84	2.71	349	**	+
F26	5.65	6.63	3.84	2.71	137	**	+
F27	8.74	6.63	3.84	2.71	122	***	+
F28	2.88	6.63	3.84	2.71	159	*	−
F29	18.16	6.63	3.84	2.71	370	***	+
F30	10.53	6.63	3.84	2.71	70	***	+
F31	1.22	6.63	3.84	2.71	54		NA

(Continues)

TABLE A2 (Continued)

Forecast	Test statistic	Critical value 99%	Critical value 95%	Critical value 90%	Obs	Significance level	Result
F32	0.74	6.63	3.84	2.71	363		NA
F33	0.00	6.63	3.84	2.71	88		NA
F35	3.17	6.63	3.84	2.71	120	*	+
F37	35.70	6.63	3.84	2.71	365	***	+
F38	15.81	6.63	3.84	2.71	363	***	+
F39	9.08	6.63	3.84	2.71	143	***	+
F40	4.02	6.63	3.84	2.71	423	**	+
F41	2.25	6.63	3.84	2.71	68		NA
F43	0.01	6.63	3.84	2.71	356		NA
F44	3.25	6.63	3.84	2.71	68	*	−
F45	1.38	6.63	3.84	2.71	260		NA
F46	2.39	6.63	3.84	2.71	334		NA
F47	6.79	6.63	3.84	2.71	317	***	+
F48	25.40	6.63	3.84	2.71	110	***	+
F49	17.61	6.63	3.84	2.71	220	***	+
F50	0.88	6.63	3.84	2.71	42		NA
F53	0.33	6.63	3.84	2.71	29		NA
F54	1.83	6.63	3.84	2.71	36		NA
F51	2.24	6.63	3.84	2.71	45		NA
F52	6.99	6.63	3.84	2.71	44	***	+

Note. + = better than a random forecast; − = worse than random forecast; NA = no statistically significant difference from a random forecast; WTI = Western Texas Intermediate.

Significance levels:

*90%, **95%, ***99%.

TABLE A3 Results test for unbiasedness Brent

Forecast	α	Se α	β	Se β	Test statistic	*F*-test critical value	Obs	Result	Significance level
MEDIAN	−0.84	1.12	0.48	0.12	18.09	3.86	468	−	
F1	−9.43	1.41	0.17	0.13	67.25	3.90	163	−	
F2	−3.54	1.19	0.28	0.15	36.74	3.87	316	−	
F3	−0.83	1.59	0.29	0.10	56.29	3.88	238	−	
F4	3.33	1.28	0.16	0.11	99.74	3.88	275	−	
F5	11.03	1.65	0.61	0.14	92.58	3.88	248	−	
F6	−3.46	1.33	0.02	0.13	78.65	3.87	361	−	
F7	14.76	1.68	0.34	0.26	87.98	4.01	60	−	
F8	7.89	2.29	0.94	0.13	14.13	3.90	156	−	
F9	−0.56	2.27	0.31	0.17	17.12	3.89	186	−	
F10	−1.94	1.22	−0.02	0.11	92.04	3.86	415	−	
F11	−2.65	1.32	−0.02	0.11	85.20	3.88	270	−	
F12	−3.34	1.19	0.19	0.11	65.70	3.89	197	−	
F13	−3.83	1.15	0.55	0.11	36.04	3.87	307	−	
F14	2.00	1.92	0.41	0.12	25.60	3.89	215	−	
F15	3.09	1.58	0.22	0.14	42.87	3.88	269	−	
F16	7.77	2.76	0.73	0.22	17.03	3.91	132	−	
F17	3.42	1.47	0.68	0.11	14.37	3.88	249	−	
F18	−1.33	2.53	−0.13	0.20	43.57	3.92	128	−	
F19	−0.02	1.36	0.57	0.14	11.30	3.87	353	−	
F20	5.11	2.02	−0.06	0.19	39.23	3.89	205	−	
F21	6.50	1.36	0.84	0.10	28.38	3.87	295	−	

(Continues)

TABLE A3 (Continued)

Forecast	α	Se α	β	Se β	Test statistic	*F*-test critical value	Obs	Result	Significance level
F22	2.70	1.41	0.61	0.18	10.49	3.87	329	–	
F23	−0.14	1.30	−0.31	0.21	40.07	3.87	363	–	
F24	2.34	1.63	−0.14	0.12	91.85	3.88	262	–	
F25	−1.39	1.31	0.06	0.09	134.04	3.87	337	–	
F26	−1.78	1.57	0.51	0.16	9.82	3.91	137	–	
F27	−0.86	1.17	0.47	0.13	17.30	3.86	415	–	
F28	4.73	1.22	0.28	0.10	57.29	3.87	327	–	
F29	3.10	1.26	0.32	0.12	50.50	3.87	371	–	
F30	6.26	4.06	1.59	0.25	6.24	3.98	70	–	
F31	−5.70	1.77	0.20	0.23	15.72	4.04	50	–	
F32	0.72	1.29	0.20	0.12	46.05	3.87	360	–	
F33	−2.75	3.66	−0.28	0.37	15.69	3.95	87	–	
F34	−6.22	2.64	−0.41	0.20	72.32	4.03	52	–	
F35	2.79	2.99	0.43	0.21	10.24	3.92	127	–	
F36	−1.01	1.19	0.17	0.11	61.96	3.86	433	–	
F37	−1.19	1.16	0.43	0.09	39.91	3.86	421	–	
F38	2.44	1.21	0.72	0.12	11.22	3.87	383	–	
F39	16.24	3.01	1.64	0.21	29.50	3.91	143	–	
F40	−0.61	1.17	0.41	0.10	33.13	3.86	419	–	
F41	−3.27	1.25	0.83	0.17	8.42	3.99	68	–	
F42	−6.11	2.37	0.67	0.35	6.66	4.08	43	–	
F43	−2.18	1.29	0.16	0.11	63.05	3.87	356	–	
F44	−13.06	2.27	0.69	0.35	33.92	3.98	70	–	
F45	−4.64	1.53	−0.19	0.10	198.81	3.88	259	–	
F46	−2.56	1.44	0.00	0.12	67.72	3.87	345	–	
F47	−2.72	1.33	0.66	0.15	11.97	3.87	325	–	
F48	1.46	2.11	0.65	0.18	19.33	3.93	110	–	
F49	10.21	1.77	1.11	0.15	37.37	3.88	240	–	
F50	−0.08	6.82	−1.03	0.44	41.14	4.01	59	–	
F51	17.78	5.48	0.47	0.49	29.82	4.07	45	–	
F52	12.49	2.72	−0.39	0.43	26.62	4.07	44	–	

Note. – = Biased forecasts; + = unbiased forecasts.

Significance levels:

*90%, **95%, ***99%.

TABLE A4 Results test for unbiasedness WTI

Forecast	α	Se α	β	Se β	Test statistic	*F*-test critical value	Obs	Result	Significance level
MEDIAN	−0.72	1.08	0.81	0.10	4.29	3.86	468	–	
F1	−6.67	1.30	0.39	0.14	46.88	3.90	163	–	
F2	−2.10	1.18	0.54	0.14	17.58	3.87	324	–	
F3	−2.13	1.70	0.53	0.10	41.05	3.88	234	–	
F4	8.09	1.35	0.35	0.14	70.18	3.90	173	–	
F5	13.63	2.10	0.59	0.15	73.94	3.90	160	–	
F6	−2.89	1.37	0.32	0.10	68.08	3.87	365	–	
F7	12.36	1.77	0.94	0.14	54.71	4.01	60	–	
F8	8.67	2.30	0.93	0.13	16.61	3.90	156	–	
F9	0.65	2.37	0.47	0.18	10.75	3.89	178	–	
F10	−2.64	1.30	0.36	0.10	53.30	3.87	355	–	

(Continues)

TABLE A4 (Continued)

Forecast	α	Se α	β	Se β	Test statistic	*F*-test critical value	Obs	Result	Significance level
F11	0.88	1.10	0.52	0.09	31.04	3.87	294	–	
F12	−2.07	1.21	0.41	0.10	48.69	3.89	193	–	
F13	−8.00	1.32	0.85	0.10	67.91	3.88	265	–	
F14	1.74	1.88	0.62	0.11	12.70	3.89	215	–	
F15	3.79	1.49	0.50	0.12	26.64	3.88	276	–	
F16	8.42	2.65	0.82	0.20	15.90	3.91	132	–	
F17	3.05	1.46	0.86	0.10	5.40	3.88	242	–	
F18	0.01	2.65	−0.05	0.19	42.93	3.91	133	–	
F19	−1.08	2.09	0.52	0.33	2.07	4.07	44	+	**
F21	7.55	1.61	0.86	0.10	24.31	3.88	235	–	
F22	2.58	1.32	1.02	0.13	3.88	3.87	321	–	
F24	4.60	1.72	0.31	0.11	50.83	3.88	246	–	
F25	1.60	1.12	0.46	0.08	52.10	3.87	349	–	
F26	−3.75	1.31	0.88	0.12	10.69	3.91	137	–	
F27	0.47	2.88	0.79	0.21	1.02	3.92	122	+	**
F28	9.60	1.41	−0.58	0.16	120.54	3.90	159	–	
F29	2.83	1.23	0.59	0.11	17.11	3.87	370	–	
F30	8.89	4.11	1.55	0.23	7.38	3.98	70	–	
F31	−3.29	1.91	0.47	0.25	18.73	4.03	54	–	
F32	1.34	1.26	0.51	0.11	20.66	3.87	363	–	
F33	0.33	3.47	−0.04	0.38	10.74	3.95	88	–	
F35	2.79	2.99	0.68	0.21	3.79	3.92	120	+	**
F37	−4.24	1.07	0.66	0.08	41.50	3.87	365	–	
F38	2.09	1.19	0.97	0.11	3.16	3.87	363	+	**
F39	16.61	2.85	1.61	0.20	34.23	3.91	143	–	
F40	−0.62	1.17	0.62	0.09	23.01	3.86	423	–	
F41	0.92	1.36	0.46	0.18	9.44	3.99	68	–	
F43	−1.06	1.29	0.42	0.10	39.78	3.87	356	–	
F44	−10.00	2.28	0.79	0.36	19.67	3.99	68	–	
F45	−3.40	1.64	0.04	0.10	142.35	3.88	260	–	
F46	−0.29	1.37	0.33	0.12	32.37	3.87	334	–	
F47	−2.15	1.35	0.82	0.15	6.19	3.87	317	–	
F48	−2.07	1.05	0.93	0.13	4.02	3.93	110	–	
F49	7.48	1.63	1.12	0.12	21.37	3.88	220	–	
F50	−16.00	7.51	−0.81	0.44	58.76	4.08	42	–	
F53	−6.93	2.68	1.84	0.44	6.73	4.21	29	–	
F54	3.45	2.28	0.73	0.35	2.31	4.13	36	+	**
F51	19.46	5.16	0.50	0.43	34.58	4.07	45	–	
F52	9.12	2.39	1.05	0.15	20.03	4.07	44	–	

Note. – = Biased forecasts; + = unbiased forecasts; WTI = Western Texas Intermediate.

Significance levels:

*90%, **95%, ***99%.

TABLE A5 Results fest for forecast encompassing Brent

Forecast	β_0	Se β_0	*p* value β_0	β_1	Se β_1	*p* value β_1	Obs	Result	Significance level
MEDIAN	0.71	0.11	.00	−0.18	0.13	.18	468	–	
F1	0.83	0.13	.00	0.15	0.16	.36	163	–	
F2	0.85	0.14	.00	−0.19	0.16	.23	316	–	
F3	0.64	0.09	.00	−0.13	0.11	.23	238	–	

(Continues)

TABLE A5 (Continued)

Forecast	β_0	Se β_0	*p* value β_0	β_1	Se β_1	*p* value β_1	Obs	Result	Significance level
F4	0.66	0.10	.00	−0.07	0.10	.48	275	−	
F5	0.12	0.12	.35	0.39	0.12	.00	248	+	***
F6	0.79	0.12	.00	−0.24	0.12	.05	361	+	***
F7	0.60	0.26	.03	0.64	0.37	.09	60	+	*
F8	0.22	0.11	.05	−0.33	0.18	.08	156	+	*
F9	0.72	0.14	.00	−0.46	0.16	.00	186	+	***
F10	0.85	0.10	.00	−0.31	0.10	.00	415	+	***
F11	1.10	0.11	.00	−0.41	0.13	.00	270	+	***
F12	0.74	0.11	.00	0.10	0.11	.36	197	−	
F13	0.51	0.11	.00	0.28	0.13	.03	307	+	***
F14	0.60	0.10	.00	−0.47	0.15	.00	215	+	***
F15	0.50	0.12	.00	−0.05	0.12	.68	269	−	
F16	0.73	0.20	.00	−0.86	0.31	.01	132	+	***
F17	0.34	0.10	.00	0.23	0.12	.06	249	+	*
F18	0.85	0.17	.00	−0.45	0.17	.01	128	+	***
F19	0.73	0.13	.00	−0.24	0.15	.12	353	−	
F20	0.65	0.15	.00	−0.47	0.15	.00	205	+	***
F21	0.13	0.10	.17	0.40	0.11	.00	295	+	***
F22	0.23	0.16	.15	0.23	0.16	.16	329	−	
F23	1.26	0.18	.00	−0.77	0.19	.00	363	+	***
F24	0.99	0.10	.00	−0.78	0.11	.00	262	+	***
F25	0.98	0.08	.00	−0.40	0.10	.00	337	+	***
F26	0.30	0.18	.10	0.97	0.28	.00	137	+	***
F27	0.79	0.12	.00	−0.27	0.15	.07	415	+	*
F28	0.53	0.09	.00	−0.05	0.09	.60	327	−	
F29	0.79	0.10	.00	−0.43	0.11	.00	371	+	***
F30	−0.44	0.23	.07	0.55	0.38	.15	70	−	
F31	0.79	0.23	.00	0.16	0.28	.58	50	−	
F32	0.82	0.11	.00	−0.33	0.12	.01	360	+	***
F33	−0.03	0.24	.89	0.01	0.22	.96	87	−	
F34	1.39	0.19	.00	−0.62	0.21	.01	52	+	***
F35	0.68	0.16	.00	−0.79	0.20	.00	127	+	***
F36	0.73	0.10	.00	−0.20	0.10	.05	433	+	*
F37	0.59	0.08	.00	−0.01	0.10	.91	421	−	
F38	0.03	0.10	.74	0.45	0.10	.00	383	+	***
F39	−0.13	0.20	.52	0.19	0.29	.51	143	−	
F40	0.53	0.10	.00	0.04	0.10	.73	419	−	
F41	0.03	0.19	.87	0.36	0.36	.32	68	−	
F42	0.49	0.32	.13	1.52	0.40	.00	43	+	***
F43	0.66	0.10	.00	−0.11	0.10	.30	356	−	
F44	0.38	0.39	.33	0.57	0.44	.20	70	−	
F45	1.10	0.09	.00	−0.57	0.10	.00	259	+	***
F46	1.02	0.10	.00	−0.64	0.12	.00	345	+	***
F47	0.44	0.14	.00	0.21	0.16	.19	325	−	
F48	−0.06	0.18	.73	0.01	0.21	.95	110	−	
F49	0.19	0.14	.18	0.18	0.19	.33	240	−	
F50	0.80	0.25	.00	−0.76	0.23	.00	59	+	***
F51	0.76	0.48	.12	−0.63	0.70	.37	45	−	
F52	0.65	0.30	.04	−0.46	0.29	.12	44	−	

Note. − = no statistical significant additional information from forecast; + = statistical significant additional information from forecast.

Significance levels:

*90%, **95%, ***99%.

TABLE A6 Results fest for forecast encompassing WTI

Forecast	β_0	Se β_0	*p* value β_0	β_1	Se β_1	*p* value β_1	Obs	Result	Significance level
MEDIAN	0.26	0.09	.00	0.17	0.11	.14	468	−	
F1	0.62	0.14	.00	0.35	0.16	.03	163	+	***
F2	0.60	0.13	.00	−0.03	0.16	.84	324	−	
F3	0.35	0.09	.00	0.00	0.10	.98	234	−	
F4	0.33	0.13	.01	0.16	0.13	.20	173	−	
F5	0.19	0.13	.15	0.06	0.15	.68	160	−	
F6	0.38	0.09	.00	0.04	0.09	.63	365	−	
F7	−0.05	0.15	.77	0.68	0.22	.00	60	+	***
F8	0.24	0.10	.02	−0.39	0.18	.03	156	+	***
F9	0.51	0.15	.00	−0.31	0.17	.07	178	+	*
F10	0.42	0.09	.00	−0.01	0.09	.93	355	−	
F11	0.49	0.08	.00	0.12	0.10	.25	294	−	
F12	0.46	0.10	.00	0.32	0.10	.00	193	+	***
F13	0.16	0.10	.11	0.64	0.12	.00	265	+	***
F14	0.38	0.09	.00	−0.39	0.14	.01	215	+	***
F15	0.24	0.09	.01	0.02	0.10	.87	276	−	
F16	0.71	0.19	.00	−0.87	0.31	.01	132	+	***
F17	0.09	0.09	.29	0.33	0.11	.00	242	+	***
F18	0.63	0.15	.00	−0.36	0.15	.02	133	+	***
F19	0.29	0.23	.22	−0.26	0.26	.31	44	−	
F21	0.07	0.08	.38	0.08	0.12	.51	235	−	
F22	−0.05	0.11	.68	0.34	0.12	.01	321	+	***
F24	0.48	0.09	.00	−0.33	0.10	.00	246	+	***
F25	0.51	0.08	.00	0.07	0.10	.44	349	−	
F26	0.08	0.12	.49	1.16	0.18	.00	137	+	***
F27	0.40	0.18	.03	−0.37	0.23	.10	122	−	
F28	0.62	0.18	.00	−0.22	0.14	.13	159	−	
F29	0.36	0.09	.00	−0.17	0.10	.09	370	+	*
F30	−0.38	0.22	.09	0.53	0.37	.16	70	−	
F31	0.44	0.24	.07	0.12	0.27	.65	54	−	
F32	0.49	0.09	.00	−0.14	0.11	.20	363	−	
F33	−0.19	0.24	.44	0.15	0.22	.51	88	−	
F35	0.54	0.16	.00	−0.62	0.20	.00	120	+	***
F37	0.29	0.08	.00	0.25	0.09	.01	365	+	***
F38	−0.10	0.09	.26	0.40	0.10	.00	363	+	***
F39	−0.04	0.19	.81	0.09	0.28	.75	143	−	
F40	0.27	0.08	.00	0.14	0.09	.10	423	−	
F41	0.08	0.20	.70	0.00	0.20	.98	68	−	
F43	0.30	0.09	.00	0.13	0.09	.14	356	−	
F44	0.33	0.40	.42	0.58	0.47	.22	68	−	
F45	0.81	0.09	.00	−0.41	0.10	.00	260	+	***
F46	0.64	0.10	.00	−0.41	0.11	.00	334	+	***
F47	0.13	0.13	.31	0.36	0.14	.01	317	+	***
F48	−0.21	0.12	.09	−0.19	0.22	.38	110	−	
F49	0.00	0.10	.97	0.17	0.14	.22	220	−	
F50	0.86	0.29	.00	−0.78	0.27	.01	42	+	***
F53	−0.44	0.28	.13	−0.17	0.42	.68	29	−	
F54	−0.48	0.33	.15	0.01	0.33	.97	36	−	
F51	0.73	0.42	.09	−0.57	0.62	.36	45	−	
F52	0.08	0.11	.50	−0.11	0.21	.60	44	−	

Note. − = no statistical significant additional information from forecast; + = statistical significant additional information from forecast; WTI = Western Texas Intermediate.

Significance levels:

*90%, **95%, ***99%.

Chapter 6

The emergence of the RMB: a "new normal" for China's exchange rate system?

with Christoph Wegener, Tobias Basse and Markus Spiwoks
contribution Frederik Kunze: 60%

Published as Working Paper:
Cege Discussion Papers, no. 348, University of Göttingen, May 2018. Download:
http://wwwuser.gwdg.de/ cege/Diskussionspapiere/DP348.pdf

Abstract

We investigate RMB pricing differentials for onshore and offshore trading. Testing for long memory, we find strong persistence in the pricing differential. Hence, the Chinese FX market in its bipolar structure still lacks basic conditions for perfectly integrated markets.

Keywords

CNY, CNH, RMB internationalization, market integration, emerging markets

JEL Classification

F31, F33, G18, C58

Highlights

- We examine the bipolar structure of the Chinese exchange rate system and document empirical evidence for cointegration of the CNH-CNY pricing differential.
- We show that the CNH-CNY spread contains strong persistence.
- Even by testing against a break in the persistence we find no indications that the two time series have become more integrated over time.
- We infer from these results that the RMB has not reached its *"New Normal"* yet.
- While Beijing's monetary policy makers may benefit from the special architecture of a neither fixed nor floating FX regime, we argue that it generates substantial challenges.
- The Chinese FX market may be prone to speculative attacks and the persistence of the CNH-CNY spread is an impediment for RMB internationalization.

1. Introduction and institutional background

China's exchange rate policy has been criticized as an important element of the government's protectionist measures to improve the price competitiveness of its export firms (see Frankel and Wei, 2007). Lately, this debate's focus shifted to RMB internationalization (see Batten and Szilagyi, 2016) and the question whether the RMB will succeed the Greenback as the new global anchor currency (see Ito, 2010). Furthermore, the RMB recently gained importance as an international payment currency (see Cheung and Rime, 2014; Funke et al., 2015). However, it is still debatable how far internationalization has gone and whether the RMB has reached a *"New Normal"*.

We argue that one major impediment for RMB internationalization is the bipolar structure of China's FX market caused by the imperfect integration of on- and offshore markets. We test market integration using long memory tests – if both markets are perfectly integrated the underlying exchange rates should be (i) cointegrated (see Cheung and Rime, 2014) and (ii) the differential should be integrated of order zero ($I(0)$). Efficient markets also excludes a mean reverting but strong dependent spread ingrated of order d ($I(d)$) with $0 < d < 0.5$ because in this case the on- and offshore currencies would be cointegrated but priced differently and thus, they would not be perfectly integrated.

The impossible trinity - or *trilemma of international financial economics* - is important in this context (see Bluedorn and Bowdler, 2010). This rule states that a country cannot have a fixed exchange rate, free capital movements and an independent monetary policy. Cooper (1999) argued that it is still unclear what type of currency regime is appropriate for a specific country. It is well documented in the literature that this choice has implications for the monetary policy options (see Cooper, 1999). In the case of rigidly fixed exchange rates all but one central banks participating in this currency regime loose the opportunity to freely use the tools of monetary policy. In fact, the end of the Bretton Woods system was a direct consequence of hopes in some countries that a central bank freed from the need to stabilize the fixed exchange rate would be able to fight more effectively against inflationary pressures (see Basse, 2006; Gray, 2007). Freely floating exchange rates were quite popular in the 1980s and 1990s. However, more recently Calvo and Reinhart (2002) diagnosed a *"fear of floating"*. Moreover, some observers seem to believe that there are alternatives (e.g., dirty floating or target zones) that can help to combine the advantages of freely floating and absolutely fixed exchange rates (see Cooper, 1999; Masson, 2001). With regard to China it could be argued that the middle kingdom has found its own middle way to create at least some flexibility in a system of fixed exchange rates and to mask the lack of CNY convertibility. This special regime is based on the existence of two different exchange rates for one currency.

As a matter of fact, the birth of the offshore RMB in Hong Kong is highly relevant in this context (see Cheung and Rime, 2014). The special administrative region hosts a large FX market and belongs to China, but has its own political system: one country, two systems (see Meyer and Revilla Diez, 2015). And the same holds for the RMB: it is China's unique currency but its quotations are subject to the place where they are traded. On the mainland the quotation CNY is used, CNH is used for offshore trading. USD/CNY and USD/CNH are both rates for the exchange of RMB against USD at different trading locations: one currency, two quotations (see Shu et al., 2015). Although market participants are dealing with the same currency, a significant difference between the USD/CNH and USD/CNY exchange rate is observable (see Cheung et al., 2017a,b). With perfect arbitrage processes the law of one price obviously would predict that the CNH-CNY spread ought to be zero – at least statistically.

The paper is structured as follows. Chapter 2 introduces the data, an initial data analysis and the methodology of our empirical analysis. In Chapter 3 we present the empirical results. Chapter 4 concludes the paper.

2. Data, methodology and initial empirical analysis

We examine weekly data for the CNY spot rate as well as the CNH spot rate from January 1st 2011 to February, 10th 2017 taken from Bloomberg. Our sample ranges from the early days of the RMB offshore market up to the episode of the current RMB weakness and includes the shifts in Beijing's FX policy. We investigate USD/CNY and USD/CNH exchange rates and corresponding pricing differential. Economic theory implies that in a perfectly integrated market these two exchange rates ought to be identical with a spread of zero (see Barros et al., 2016). Since the spread is time dependent (see Figure 1) we examine whether the pricing differences occur systematically (see Craig et al., 2013).

Finding cointegration among the two exchange rates would imply convergence between these financial market prices. In fact, Becker and Hall (2007) argued that cointegration is a sign for convergence among non-stationary time series. As stated above, it makes sense to assume cointegration for the two FX quotations. More specifically, market efficiency should lead to cointegration among the prices of two almost identical financial assets (e.g. Alexander, 1999; Westerlund and Narayan, 2013). Despite the fact that capital controls might cause a non-zero spread, a long-term relationship between the USD/CNY and USD/CNH exchange rate should exist.

The existence of cointegration between the two FX rates is well documented in the literature (see Cheung and Rime, 2014). Therefore, the CNH-CNY-spread should be stationary

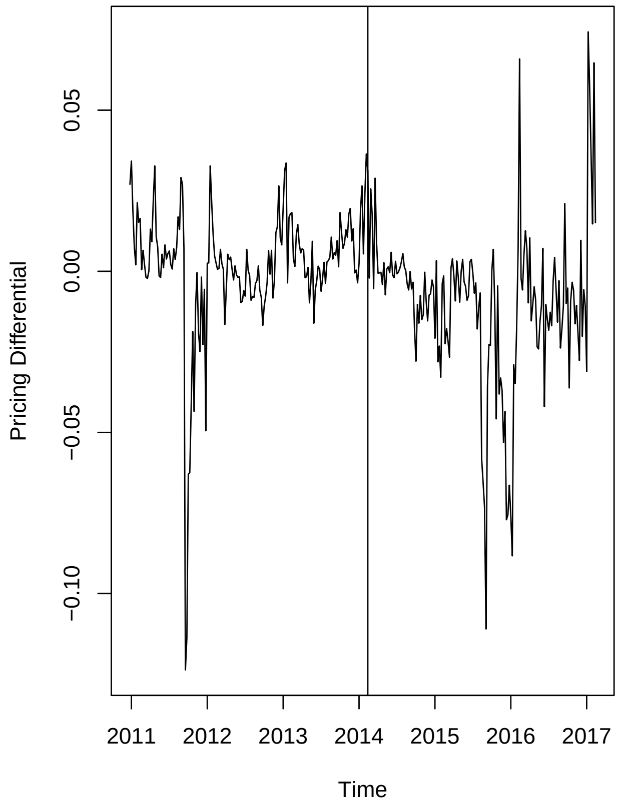


Fig. 1. The graph shows the pricing differential between the CNY and CNH quotations from January 2011 to February 2017. The vertical line indicates a potential break in the persistence on February 7th, 2014.

or integrated of order 0 ($I(d)$ with $d = 0$). Thus, we used recently developed techniques of time series analysis to produce additional evidence of relevance in this context. However, cointegration covers the case of two time series sharing a common stochastic trend and a stationary linear combination of both series ($I(d)$ with $0 \leq d < 0.5$).

In this case the spread (y_t) might also be strongly persistent (for $0 < d < 0.5$) and the assumption of perfectly integrated RMB on- and offshore markets would not be fulfilled. Thus, we estimate the degree of integration d and assume that the spread follows a fractionally integrated process of the form

$$\Phi(B)(1 - B)^d y_t = \Psi(B)\epsilon_t \tag{1}$$

where all roots of the polynomials $\Phi(B)$ and $\Psi(B)$ are assumed to lie outside the unit circle and ϵ_t is independent and identically distributed with $E(\epsilon_t) = 0$, $sup_t E(\epsilon_t^2) < \infty$.

The degree of persistence $d \in [0, 0.5)$ determines the degree of integration of the spread and $(1-B)^d$ is defined by its binomial expansion

$$(1-B)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} B^j \tag{2}$$

with $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ and B as the Backshift operator, i.e. $By_t = y_{t-1}$.

To estimate the degree of persistence d we employ the local Whittle estimator (see Shimotsu and Phillips, 2005; Shimotsu, 2010). However, we want to test the hypothesis of $H_0 : d = \hat{d}$. Thus, we employ the Augmented Lagrange Multiplier (ALM) test by Demetrescu et al. (2008). The authors suggest a lag augmented version of the Lagrange multiplier test by Robinson (1991). This procedure is based on the regression

$$y_t = \phi y^*_{t-1} + a_1 y_{t-1} + a_2 y_{t-2} + ... + a_p y_{t-p} + \varepsilon_t \text{ for } t = p+1, ..., T \tag{3}$$

with $y^*_{t-1} = \sum_{j=1}^{t-1} \frac{y_{t-j}}{j}$, p as the number of lags in the augmentation, which grows with the sample size, and ε_t as an innovation process. The authors retain limiting normality of the t_ϕ-statistic, which is used to test the null hypothesis $H_0 : \phi = d = 0$. However, we use the estimation result of the local Whittle estimator to use $\hat{d}$th differences of the spread. By using this procedure we are able to test $H_0 : d = \hat{d}$.

Furthermore, we are particularly interested in whether d remains constant over time. Thus, we use the methodology proposed by Sibbertsen and Kruse (2009) to test the hypothesis

$$H_0 : d = d_0, \ \forall \ t \text{ vs. } H_1 : \begin{cases} d = d_1 \text{ for } t = 1, ..., [\tau T] \\ d = d_2 \text{ for } t = [\tau T] + 1, ..., T \end{cases} . \tag{4}$$

Here, $[\tau T]$ denotes the biggest integer smaller than τT with τ as the relative breakpoint estimator and T as the number of observations.

The authors restricted $0 \leq d_0 < \frac{3}{2}$ under H_0 and $0 \leq d_1 < \frac{1}{2}$ and $\frac{1}{2} \leq d_2 < \frac{3}{2}$ under the alternative. Moreover, d_1 and d_2 can be exchanged, so a break from stationary to non-stationary long-memory and vice versa can be investigated. Thus, we test against a break in the persistence in the spread using the estimated d under the null hypothesis by the local Whittle estimator.

Thus, our procedure works in three steps:

1. Estimating the persistence parameter d by the local Whittle estimator. The estimated counterpart is indicated by $\hat{d}$
2. Testing the Hypothesis of $H_0 : d = 0$ and $H_0 : d = \hat{d}$ against the fractional alternatives
3. Testing against a break in the persistence

3. Empirical results

We start our empirical analysis by considering the autocorrelation function in Figure 2. The

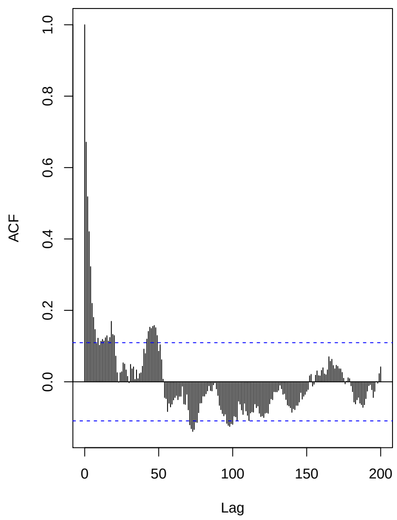


Fig. 2. The graph shows the autocorrelation function of the pricing differential for the full sample with 200 lags (approx. 4 years).

graph shows the autocorrelation of the spread. By considering the function it seems that it declines quite slowly which might be a first indication for strong dependence.

We then estimate the degree of persistence d employing the local Whittle estimator ($\hat{d}$) and draw inference about d considering the results of the test by Demetrescu et al. (2008). For the estimation of d we use two different bandwidths: $m = T^{0.55}$ ($\hat{d}_{0.55}$) and $m = T^{0.70}$ ($\hat{d}_{0.70}$). For the ALM test by Demetrescu et al. (2008), we simulate critical values for the test version with recursive de-meaning for the small sample of 320 observations. We use $100,000$ steps within the Monte Carlo simulation. See Table 1 for the results.

	$\hat{d}_{0.55}$	$\hat{d}_{0.70}$	
Whittle estimation	0.30	0.48	
	$H_0 : d = 0$	$H_0 : d = \hat{d}_{0.55}$	$H_0 : d = \hat{d}_{0.70}$
ALM	1.16*	-0.26	-1.09

Table 1: This table reports the results of the estimation of d as well as the results of the ALM test. We use simulated critical value with $100,000$ test replications for the version of the test using the recursive de-meaning. The critical value are 1.05 (1%), 1.43 (5%), 2.10 (10%).

The results support the hypothesis of a strongly dependent spread and we must reject the null hypothesis of $d = 0$ on a significance level of 10% while we cannot reject the hypotheses of $d = \hat{d}_{0.55}$ and $d = \hat{d}_{0.70}$.

Furthermore, we are interested whether the persistence is stable over time. To test against a break in the persistence we use the test proposed by Sibbertsen and Kruse (2009). All results are reported by Table 2.

τ_{low}	0.2	0.4
τ_{up}	0.8	0.6
$d_0 = \hat{d}_{0.55}$	Cannot reject H_0	Increasing Persistence
	-	February 7th, 2014
$d_0 = \hat{d}_{0.7}$	Cannot reject H_0	Increasing Persistence
	-	February 7th, 2014

Table 2: This table reports the results of the structural break test by Sibbertsen and Kruse (2009). Here, we consider two cases concerning the choice of τ_{low} and τ_{up}.

We find results in favor of a break for a combination of $\tau_{low} = 0.4$ and $\tau_{low} = 0.6$ – that means we allow in this particular example for a break from observation $0.4T$ to $0.6T$. However, if we set $\tau_{low} = 0.2$ and $\tau_{low} = 0.8$ we cannot reject the null hypothesis. Thus, even if we find indication for a break the persistence increases – this gives even more support for the finding that China's FX markets are far away from being perfectly integrated.

To summarize our empirical results: We find evidence for a strong dependent pricing differential – which does not support that the CNH and CNY quotations are integrated.

Furthermore, by investigating potential structural breaks we find no support for decreasing persistence at all – which underlines the result that the *"New Normal"* of China's FX markets is not reached. Figure 3 and Figure 4 show also indications for strong dependence before and after the break point. Our results confirm earlier findings in the literature in the sense that financial market practitioners have to consider basis risk when hedging CNY exposure with CNH contracts (see Craig et al., 2013).

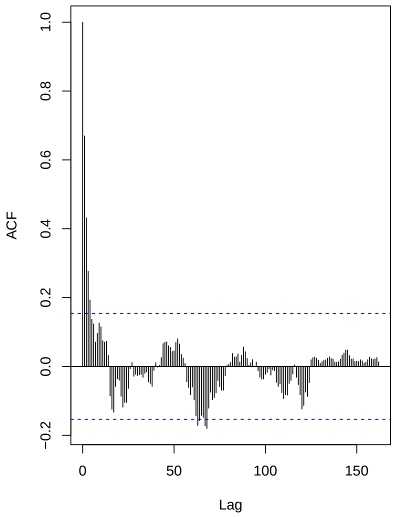


Fig. 3. The graph shows the autocorrelation function of the pricing differential until the break point (February 7th, 2014) sample with 163 lags (approx. 3 years).

4. Conclusions

We examine the CNH-CNY pricing differential and document empirical evidence for cointegration. This result can be seen as an indication for ongoing arbitrage processes between the onshore and the offshore market. However, we also show that the spread contains strong persistence. Even by testing against a break in the persistence we find no indications that the two time series have become more integrated over time. Regarding the present state of the Chinese FX market these results have to be seen as empirical evidence that the RMB has not reached its *"New Normal"* yet. Our findings show a substantial weakness of the *"one currency, two quotations approach"*. The enduring CNH-CNY pricing differential reduces the effectiveness of the offshore RMB as an hedging tool. Hence, Beijing's monetary policy makers may benefit from the special architecture of a neither fixed nor floating FX

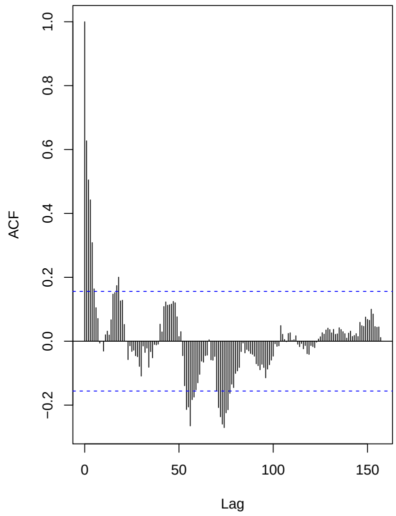


Fig. 4. The graph shows the autocorrelation function of the pricing differential after the break point (February 7th, 2014) sample with 157 lags (approx. 3 years).

regime. However, the bipolar structure of China's FX markets generates substantial challenges. Speculators, for example, have been utilizing the Hong Kong market to bet against the RMB (see Neely, 2017). Prevailing pricing differentials between the two FX rates made substantial interventions both on the mainland and in the offshore market necessary and demonstrated the costs of a managed float with two exchange rates.

References

Alexander, C. (1999). Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 357(1758):2039–2058.

Barros, C. P., Gil-Alana, L. A., and Chen, Z. (2016). Exchange rate persistence of the Chinese Yuan against the US Dollar in the NDF market. *Empirical Economics*, 51(4):1399–1414.

Basse, T. (2006). Floating exchange rates and inflation in Germany: Are external shocks really irrelevant? *Economics Letters*, 93(3):393–397.

Batten, J. A. and Szilagyi, P. G. (2016). The internationalisation of the RMB: New starts, jumps and tipping points. *Emerging Markets Review*, 28:221 – 238.

Becker, B. and Hall, S. (2007). Measuring convergence of the new member countries exchange rates to the Euro. *Journal of Financial Transformation*, 19:20–25.

Bluedorn, J. C. and Bowdler, C. (2010). The empirics of international monetary transmission: identification and the impossible trinity. *Journal of Money, Credit and Banking*, 42(4):679–713.

Calvo, G. A. and Reinhart, C. M. (2002). Fear of floating. *The Quarterly Journal of Economics*, 117(2):379–408.

Cheung, Y.-W., Hui, C.-H., and Tsang, A. (2017a). The Renminbi central parity: An empirical investigation. *BOFIT Discussion Papers No. 07/2017*.

Cheung, Y.-W., Hui, C.-H., and Tsang, A. (2017b). The RMB central parity formation mechanism after august 2015: A statistical analysis. *HKIMR Working Paper No. 06/2017*.

Cheung, Y.-W. and Rime, D. (2014). The offshore Renminbi exchange rate: Microstructure and links to the onshore market. *Journal of International Money and Finance*, 49:170–189.

Cooper, R. N. (1999). Exchange rate choices. *Sneddon Little, J. and Olivei, G. P.: Rethinking the International Monetary System*, pages 99–123.

Craig, R. S., Hua, C., Ng, P., and Yuen, R. (2013). *Development of the Renminbi Market in Hong Kong SAR: Assessing Onshore-Offshore Market Integration*. Number 13-268. International Monetary Fund.

Demetrescu, M., Kuzin, V., and Hassler, U. (2008). Long memory testing in the time domain. *Econometric Theory*, 24(01):176–215.

Frankel, J. A. and Wei, S.-J. (2007). Assessing China's exchange rate regime. *Economic Policy*, 22(51):576–627.

Funke, M., Shu, C., Cheng, X., and Eraslan, S. (2015). Assessing the CNH–CNY pricing differential: Role of fundamentals, contagion and policy. *Journal of International Money and Finance*, 59:245–262.

Gray, W. G. (2007). Floating the system: Germany, the United States, and the breakdown of Bretton Woods, 1969–1973. *Diplomatic History*, 31(2):295–323.

Ito, T. (2010). China as number one: how about the Renminbi? *Asian Economic Policy Review*, 5(2):249–276.

Masson, P. R. (2001). Exchange rate regime transitions. *Journal of Development Economics*, 64(2):571–586.

Meyer, S. and Revilla Diez, J. (2015). One country, two systems: How regional institutions shape governance modes in the greater Pearl River Delta, China. *Papers in Regional Science*, 94(4):891–900.

Neely, C. (2017). The peoples bank of china boosts the Yuan. *Economic Synopses*, (3):1–2.

Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, 47(1):67–84.

Shimotsu, K. (2010). Exact local whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory*, 26(02):501–540.

Shimotsu, K. and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics*, 33(4):1890–1933.

Shu, C., He, D., and Cheng, X. (2015). One currency, two markets: the Renminbi's growing influence in Asia-Pacific. *China Economic Review*, 33:163–178.

Sibbertsen, P. and Kruse, R. (2009). Testing for a break in persistence under long-range dependencies. *Journal of Time Series Analysis*, 30(3):263–285.

Westerlund, J. and Narayan, P. (2013). Testing the efficient market hypothesis in conditionally heteroskedastic futures markets. *Journal of Futures Markets*, 33(11):1024–1045.

Chapter 7

Predicting exchange rates in Asia: new insights on the accuracy of survey forecasts

contribution Frederik Kunze: 100%

Submitted:
Journal of Forecasting

Also published as Working Paper:
Cege Discussion Papers, no. 326, University of Göttingen, November 2017. Download: http://wwwuser.gwdg.de/ cege/Diskussionspapiere/DP326.pdf

Abstract

This paper evaluates aggregated survey forecasts with forecast horizons of 3, 12, and 24 months for the exchange rates of the Chinese yuan, the Hong Kong dollar, the Japanese yen, and the Singapore dollar vis-à-vis the US dollar using common forecast accuracy measures. Additionally, the rationality of the exchange rate predictions is assessed utilizing tests for unbiasedness and efficiency. All investigated forecasts are irrational in the sense that the predictions are biased. However, these results are inconsistent with an alternative measure of rationality based on methods of applied time series analysis. Investigating the order of integration of the time series and using cointegration analysis, empirical evidence supports the conclusion that the majority of forecasts are rational. Regarding forerunning properties of the predictions, the results are less convincing, with shorter term forecasts for the tightly managed USD/CNY FX regime being one exception. As one important evaluation result, it can be concluded, that the currency regime matters for the quality of exchange rate forecasts.

Keywords

Exchange rates, survey forecasts, forecast evaluation, forecast acccuracy, forecast rationality, cointegration, impulse response analysis

JEL Classification

F31, F37, G17, O24

Highlights

- I investigate monthly survey based exchange rate forecasts for the exchange rates of the Chinese yuan, the Hong Kong dollar, the Japanese yen as well as the Singapore dollar vis-à-vis the US dollar.
- I find strong empirical evidence for topically oriented trend adjustment (TOTA) forecasting behavior for all four exchange rates.
- Especially for shorter forecast horizons the naïve *no change* prediction seems to outperform the survey forecasts with the three months forecasts for the USD/CNY exchange rate being the only exception.
- I show that for all four currency regimes the forecasts and actual exchange rates share the same order of integration and find empirical evidence for long term relationships.
- Impulse response analysis documents that for all currency regimes under investigation only the three months forecast for the managed USD/CNY exchange rate has forerunning properties.
- I present empirical evidence indicating that the forecasts for the managed Chinese exchange rate systems are closest to be being rational and forerunning.
- Further research should focus on the predictability of regime shifts in managed exchange rate systems and currency boards as well as the credibility of fixed exchange rate regimes and its impact on forecast accuracy.

1. Introduction

The foreign exchange (FX) market belongs to the largest financial markets globally (see, for example, Sosvilla-Rivero and Ramos-Herrera, 2013). Additionally, exchange rates are crucial price variables in open economies and international finance (see Dreger and Stadtmann, 2008; Dick, MacDonald, and Menkhoff, 2015) and have a significant impact on foreign trade and cross border investments (see, for example, Kan, 2017). Hence, the understanding of the price building processes in the FX market is relevant for decision makers in general and for forecasters in particular (Ince and Molodtsova, 2017). This is also true for Asian currencies, because with respect to global GDP growth, cross border investments and world trade Asian economies are becoming increasingly important, and, in recent years, asset managers outside Asia seeking to diversify their investment positions are focusing on Asia's financial markets (see, for example, Dunis and Shannon, 2005).

During the global financial crisis, various currencies have been under pressure of a pronounced and unanticipated appreciation of the US dollar (Fratzscher, 2009). These movements are a potential source of severe economic and financial consequences, because uncertainty regarding exchange rate movements might hinder economic activity including cross border trade (see, for example, Thorbecke, 2008; Chit, Rizov, and Willenbockel, 2010). Hayakawa and Kimura (2009) have shown that especially in East Asia intra-regional trade is negatively affected by FX volatility. In addition to that, both foreign and domestic companies have to deal with currency risk (see Aggarwal, Chen, and Yur-Austin, 2011; Aggarwal, 2013). De Grauwe and Markiewicz (2013) note that especially in non-U.S. equity markets, and in certain sub-periods, currency risks have been the major driver of risk premiums in the stock market. Since currency risk is non-neglectable, FX forecasts may create economic value for financial market participants, central banks, policy decision makers as well as importers and exporters. Following Duffy and Giddy (1975), FX forecasts have to be seen as *"significant inputs to decisions concerning practically every aspect of international business"* (see Duffy and Giddy, 1975, p. 1).

Owing to the growing importance of Asia's major economies and the region's most important financial centers, this paper focuses on the investigation of forecasts for exchange rates of the Chinese yuan, the Hong Kong dollar, the Japanese yen, and the Singapore dollar vis-à-vis the US dollar. Using measures of forecast accuracy, common tests for rationality (i.e. unbiasedness and efficiency), as well methods of applied time series analysis (i.e. cointegration analysis and impulse response functions), the quality of survey forecasts for these exchange rates will be assessed. As a consequence of significant differences in FX regimes in the currency areas under investigation, it will also be examined whether there exists empirical evidence for regime dependent variations in the quality of survey forecasts.

Given the relevance of the FX regime on the price building processes for exchange rates, it is expected that survey forecasts do vary regarding accuracy and rationality. For forecasters classifying the relevant exchange rate regime is a crucial step (see, for example, Von Spreckelsen, Kunze, Windels, and von Mettenheim, 2014). Against the background of a strong influence of policy makers on exchange rates in managed FX regimes and the statutory provisions in currency boards it is standing to reason, that fixed and strongly managed exchange rates are "easier" to forecast, because policy makers are following an anticipated or even mandatory agenda. On the other hand, free float FX regimes seem to lack these policy guidelines which makes it more difficult to forecast the future path of exchange rates. Hence, while focusing on the investigation of survey based exchange rate forecasts provided by *Consensus Economics* for four different currency regimes, this paper seeks to combine two strands of research dealing with FX forecast evaluation and FX regimes and, while doing so, to fill a relevant gap in the literature.

The remainder of the paper is structured as follows. Chapter 2 delivers an overview of the relevant literature. Chapter 3 lists the chosen currency areas as well as the data set and describes the underlying regimes. Before exhibiting the methodological framework Chapter 4 delivers initital empirical results. Chapter 5 presents the evaluation results. Chapter 6 discusses the implications of the results in more detail and concludes the paper.

2. Literature overview

Frenkel and Mussa (1980) note that policy makers have to deal with FX market fluctuations and that the predictability of FX rates is important in this context. Nevertheless, they also comment that *"the facts indicate, however, that exchange-rate changes are largely unpredictable"* (see Frenkel and Mussa, 1980, p. 374). In their seminal paper Meese and Rogoff (1983) discuss the inability of fundamental models to forecast exchange rates in detail and motivate a large body of research dealing with the quality of FX predictions (see, for example, Frankel and Rose, 1994; Cheung and Chinn, 1998; Kilian and Taylor, 2003; Frenkel, Mauch, Rülke, et al., 2017, to name but a few). Despite the findings of Meese and Rogoff (1983) as well as Frenkel and Mussa (1980), FX survey forecasts are still of relevance for economic agents. Especially for financial market practitioners and decision makers, unable or unwilling to build up forecasting models for numerous financial variables of their own, survey forecasts (for exchange rates) may serve as one alternative. Ter Ellen, Verschoor, and Zwinkels (2013) annotate that investors in the FX market have to gather costly information to be able to form expectations. Not surprisingly, FX survey forecasts (both in aggregated and disaggregated form) have been intensively assessed with regard to their rationality (as regards, for example, unbiasedness and efficiency) as well as accuracy.

In the context of the evaluation of FX survey forecasts, data sets from numerous providers have been used.[1] Especially, the rationality of FX predictions, which is also in the focus of this paper, has been analyzed intensively (see, for example, Audretsch and Stadtmann, 2005; De Grauwe and Markiewicz, 2013; Rülke and Pierdzioch, 2013) with diverging results. Audretsch and Stadtmann (2005), who focus on disaggregated survey forecasts from the *Wall Street Journal*, find no evidence for the assumption of rational agents forming homogeneous expectations. Dominguez (1986) investigates survey forecasts for FX markets in emerging economies and concludes that the rational expectation hypothesis has to be rejected. Avraham, Ungar, and Zilberfarb (1987) find similar evidence for the Israeli shekel. Frenkel, Mauch, Rülke, et al. (2017) examine forecasters' rationality regarding exchange rate predictions. Their work is of high relevance for the focus of this paper because the authors empirically assess differences between currency forecasts for developed and emerging economies. They present a link between forecast accuracy as well as rationality and different currency areas and, hence, fill a relevant gap in the literature on FX survey forecast evaluation.

Focusing on a specific Asian market, Tsuchiya and Suehara (2015) investigate forecasts provided by *Consensus Economics* for the USD/CNY exchange rates between July 2005 and December 2012. The authors examine the directional accuracy of the forecasts and conclude that forecasters are not inferior to a naïve benchmark. Tsuchiya and Suehara (2015) find evidence for changes in FX forecasters herding behavior due to the financial crisis and central bank interventions and conclude that the monetary policy respectively the currency regime is of high relevance when evaluating FX forecasts. Duffy and Giddy (1975) compare the predictability of flexible and fixed exchange regimes and state that predictions for flexible exchange rates are futile. The monetary authorities' influence on the Japanese FX market (i.e. the exchange rate between the Japanese yen and the US dollar) and the resulting forecast dispersion has been assessed by Reitz, Stadtmann, and Taylor (2010). Also using data provided by *Consensus Economics* the authors conclude that while with increasing volatility of the USD/JPY FX forecast dispersion increases, policy interventions in the FX market have dampening effects on forecast dispersion. MacDonald and Nagayasu (2015) evaluate USD/JPY forecasts provided by the *Japan*

[1] In addition to the forecasts provided by *Consensus Economics* (Leitner and Schmidt, 2006; Beine, Bénassy-Quéré, and MacDonald, 2007; Jongen, Verschoor, Wolff, and Zwinkels, 2012; Ince and Molodtsova, 2017) predictions collected via the *Wall Street Journal (WSJ) Poll* (Audretsch and Stadtmann, 2005; Mitchell and Pearce, 2007; Frenkel, Rülke, and Stadtmann, 2009; Rülke, Frenkel, and Stadtmann, 2010), *FX Week* (e.g. Ter Ellen et al., 2013), *Bloomberg* (e.g. Pancotto, Pericoli, and Pistagnesi, 2014), *Forecasts Unlimited* (Bacchetta, Mertens, and Van Wincoop, 2009; Beckmann and Czudaj, 2017a,b; Ince and Molodtsova, 2017), *Blue Chip Forecasts* (Baghestani, 2010), *Reuters* (e.g. Bofinger and Schmidt, 2003) as well as the *ZEW Finanzmarkttest* (e.g. Bofinger and Schmidt, 2003; Leitner and Schmidt, 2006; Spiwoks and Hein, 2007; Heiden, Klein, and Zwergel, 2013) have been investigated. The data sets provided from *Consensus Economics* which come to use in this paper carry the advantages of a large historical database and three forecast horizons.

Center for International Finance (JCIF) survey and conclude that predictions are irrational.

When examining Asia's foreign exchange markets it has to be acknowledged that, unlike in the EMU, there does not exist a common currency area in Asia. Even more than that, currency regimes do vary significantly. Following, for example, Klein and Shambaugh (2008) it can be stated that the choice of the FX regime does matter and is a central topic in international finance. For a lot of currency areas, the bipolar segmentation between hard pegs or free floating only appears on the surface (Obstfeld and Rogoff, 1995; Calvo and Reinhart, 2002). For example, Moosa and Li (2017) remark that in the case of China the identification of the FX regime is important for the debate whether the Chinese currency is undervalued or not. This, in turn, is a relevant decision variable for forecasters. Having said that, it has to be taken into consideration that with the course of time monetary respectively FX regimes have changed manifold (Hernandez and Montiel, 2003; Bordo, Choudhri, Fazio, and MacDonald, 2017). Moreover, most emerging economies do not have a long track record when it comes to floating FX regimes (Kohlscheen, 2014). As a matter of fact, professional forecasters might also adjust their behavior when publishing forecasts for different FX regimes. Chinn and Frankel (1994) note that forecasters might be reluctant to deliver naïve predictions – especially for unstable currencies. Duffy and Giddy (1975), focusing on flexible exchange rates, present results indicating that for major exchange rates forecasting is not profitable.

3. Data set and FX regime classification

The monthly mean of survey FX forecasts with regard to four Asian currencies collected by *Consensus Economics* is used. The data set contains three different forecast horizons (3, 12, and 24 months) and ranges from January 1999 to March 2017. Forecasts for the following exchange rates will be evaluated: Chinese yuan against the US dollar (USD/CNY), Hong Kong dollar against the US dollar (USD/HKD), Japanese yen against the US dollar (USD/JPY), and Singapore dollar against US dollar (USD/SGD).[2]

The four exchange rates are suitable for the purpose of this paper for various reasons. Firstly, the currency areas are of high economic relevance in Asia and / or fulfill an important role in global financial markets. In terms of nominal GDP, for example, China respectively Japan are the two largest economies in Asia and number two respectively three globally. Furthermore, China is the world's largest trade nation (as measured by exports plus imports) and the Japanese currency plays an important role as a safe haven

[2]Throughout the paper all exchange rates are given as units per US dollar. Hence, a rise in the exchange rate corresponds to a US dollar appreciation and a lower exchange rate corresponds to a depreciation of the US dollar.

asset for investors. Hong Kong and Singapore as city states share the common characteristic to belong to the world's most sophisticated financial centers (see, for example, Tse and Yip, 2006; Woo, 2016). Secondly, the four currency areas do vary significantly when it comes to their FX regimes (see, for example, Tse and Yip, 2006; Cheung, Chinn, and Fujii, 2007; Chow, 2007; Takagi, 2007). Since this study also aims to examine possible differences in the forecast accuracy controlling for the FX regime this classification is important for the purpose of this paper.

However, the officially announced *de jure* exchange rate regimes do very often deviate from the observable *de facto* exchange rate regimes (see Obstfeld and Rogoff, 1995; Calvo and Reinhart, 2002; Klein and Shambaugh, 2008; Patnaik, Shah, Sethy, and Balasubramaniam, 2011). Additionally, currency regimes are not necessarily stable over time. Since 1998, the IMF publishes *de facto* classifications of the countries' FX regimes (see Kokenyne, Veyrune, Habermeier, and Anderson, 2009). To ensure comparability the FX regime classification in this paper is taken from the official IMF publication *Annual Report on Exchange Arrangements and Exchange Restrictions* (i.e. years 2000 to 2016). This *de facto* classification is useful because it is regularly updated (Patnaik, Shah, Sethy, and Balasubramaniam, 2011). Table 1 summarizes the data used. Following this classification the "degree of control" ranges from the USD/JPY (least) to the USD/HKD (most). Furthermore, the USD/CNY FX rate is more strongly controlled than the USD/SGD FX rate.

Table 1: Overview of monthly FX forecasts under investigation

Country	Start	End	Quotation	Current FX regime
PR China	01/1999	03/2017	USD/CNY	Other Managed Arrangement
Hong Kong	01/1999	03/2017	USD/HKD	Currency Board
Japan	01/1999	03/2017	USD/JPY	Free Float
Singapore	01/1999	03/2017	USD/SGD	Stabilized Arrangement

Sources: IMF AREAR 2016, *Consensus Economics*

Given the data set examined in this paper and the potential for regime shifts some further preliminary thoughts regarding the definition of the four currency areas are warranted.[3]

In the current political and academic debate China's exchange rate regime receives a lot of attention. Here, the focus lies on the People's Bank of China's influence on the exchange rate vis-à-vis the US dollar or a basket of currencies. (Kan, 2017, delivers

[3] Following the reasoning of Hernandez and Montiel (2003); Bordo, Choudhri, Fazio, and MacDonald (2017) it has to be accounted for possible shifts in the FX regimes investigated here.

a good summary of the discussion). In recent years, the Chinese authorities carried out far reaching adjustments to the FX regime (Cheung, Hui, and Tsang, 2016, 2017). Most importantly, in 2005, the central bank announced that the Chinese currency would switch to a managed float regime *"with reference to a basket of currencies"* (see Tian and Chen, 2013, p. 16). Moosa and Li (2017) deliver a clear overview of the 2005 adjustments and the subsequent steps. The IMF's *de facto* classification, as well as important announcements by the Chinese government, can be found in Table 12 in the Appendix.[4]

In contrast to the Chinese FX regime the Japanese currency's floating exchange rate regime is already in place since 1973 (see, for example, Hamada and Hayashi, 1985; Hutchison and Walsh, 1992, as well as Table 13 in the Appendix). Hence, the existence of regime shifts with relevance for the focus of this paper can be ruled out completely. The same holds true for the currency board system of Hong Kong, which has been established as early as 1983 and, since, has been in place (see, for example, Ho, 2002; Cook and Yetman, 2014, as well as Table 14 in the Appendix). In contrast, the Monetary Authority of Singapore (MAS) operates an exchange rate based monetary policy and, hence, the FX rate is practically the instrument to steer both output and inflation (see, for example, Siregar, Har, et al., 2001; Devereux, 2003; Chow, 2007; Chow, Lim, and McNelis, 2014, as well as Table 15 in the Appendix). As a consequence there does not exist a free floating USD/SGD exchange rate, but a managed regime which occasionally results in the more volatile exchange rate vis-à-vis the US dollar in comparison to the Hong Kong dollar (Devereux, 2003).[5] This regime lasts for the entire time period under investigation.

4. Initial empirical analysis and methodological framework

Before starting to evaluate the forecasts for the FX rates it is necessary to investigate the trending behavior of the exchange rate time series and the corresponding forecasts. To test for unit roots, the non-parametric test procedure proposed by Phillips and Perron (1988) will be used. The results of the PP (Phillips and Perron) unit root tests in levels and first differences (Δ) are given in Table 2 below.

[4]The regime change in 2005 might also be relevant for the focus of this paper. Because of that, robustness checks will be executed. It is important to note that for the USD/CNY FX rate controlling for the notable regime shift in July 2005 would lead to a shorter data period ranging from August 2005 to March 2017.

[5]Based on standard deviation of log differences as an approximation of returns the volatility is highest for the USD/JPY, followed by the USD/SGD, USD/CNY, and USD/HKD.

Table 2: Phillips & Perron (PP) unit root tests

Country	Time Series	p-value (pv) Level	Δ
USD/CNY	3M forecast	0.99	0.01
	12M forecast	0.99	0.01
	24M forecast	0.99	0.01
	actual	0.99	0.01
USD/HKD	3M forecast	0.03	0.01
	12M forecast	0.01	0.01
	24M forecast	0.01	0.01
	actual	0.04	0.01
USD/JPY	3M forecast	0.63	0.01
	12M forecast	0.77	0.01
	24M forecast	0.79	0.01
	actual	0.62	0.01
USD/SGD	3M forecast	0.61	0.01
	12M forecast	0.76	0.01
	24M forecast	0.85	0.01
	actual	0.54	0.01

Reported are the results (p-values) of the PP unit roots tests for the forecasts and the actual exchange rate for USD/CNY, USD/HKD, USD/JPY, and USD/SGD with the forecast horizons of 3, 12, and 24 months in levels and first differences (Δ).

Results unit root tests Following Table 2 above, with exception of the actual Hong Kong dollar exchange rate and the corresponding forecasts statistical evidence for unit roots are found for all time series under investigation.[6] These results will be discussed in more detail below.

[6]The results for the USD/CNY FX rate presented in Table 2 range from 01/1999 to 02/2017. The results for time period 08/2005 to 03/2017 are in general similar and will not be reported in order to preserve space.

4.1. Measures of forecast accuracy

Theil's U measure In a first step, straightforward measures of forecast evaluation will be used. It will be started by calculating the Theil's U measure (Theil, 1955). Equation 1 below shows the Theil's U measure.

$$U = \frac{RMSE^{Forecast}}{RMSE^{Naive}} \tag{1}$$

The major advantages of this metric, which compares two competing forecasts, are that it is easy to calculate, uncomplicated to interpret as well as dimensionless. Recently, Ahmed, Liu, and Valente (2016) as well as Byrne, Korobilis, and Ribeiro (2016) utilize the Theil's U measure in the context of evaluating FX forecasts. The Theil's U can easily be calculated as the fraction of the root mean squared error (RMSE)[7] of the forecast and the naïve prediction.[8]

Diebold Mariano test Additionally, the Diebold Mariano (DM) test to check for equal predictive ability of the survey forecast and the naïve prediction will be applied (see also Diebold and Mariano, 1995). In difference to the Theil's U measure the DM test allows to statistically validate possible differences in forecast accuracy between the survey forecasts and benchmark measures like the naïve prediction (see, for example, Kunze, Wegener, Bizer, and Spiwoks, 2017).

TOTA coefficient In the context of forecast evaluation, the TOTA coefficient developed by Andres and Spiwoks (1999) is a reasonable supplement to the Theil's U measure and the DM test. The TOTA coefficient, as the fraction of the coefficient of determination of the FX forecast and the actual FX rate at $t+h$ ($R^2_{Forecast;Actual,t+h}$) as well as the FX forecast and the actual FX rate at t ($R^2_{Forecast;Actual,t}$), is shown in the Equation 2 below.

$$TOTA = \frac{R^2_{Forecast;Actual,t+h}}{R^2_{Forecast;Actual,t}} \tag{2}$$

A comparably stronger linear relationship between the forecast and the FX rate at t results in a TOTA coefficient smaller than 1. On the other hand, a TOTA coefficient larger than 1 can be seen as an indication for the usefulness of the forecasts by definition (see Andres and Spiwoks, 1999). As stated above, this procedure has already been

[7]The RMSE itself is an often applied measure of forecast accuracy (see also Leitch and Tanner, 1991; Hyndman and Koehler, 2006; Herwartz and Schlüter, 2017; Ryan and Whiting, 2017).

[8]The naïve prediction is used as competing forecast for the survey prediction. Throughout this paper the naïve prediction is defined as the *no change forecast*, (i.e. $\hat{S}_{t+h} = S_t$) whereas S_t is the FX spot rate at t and $\hat{S}_{t+h}$ is the forecast of the FX spot rate for $t+h$. Alternatively, instead of the *no change forecast* the naïve prediction could have also been defined as *trend following*. However, given the specific data set (e.g. turning points respectively trend reversals) it is presumed that the *no change forecast* to be more appropriate for the purpose of this paper.

applied to FX forecasts. For example, Baghestani (2010) concludes that forecasts for the trade weighted USD/EUR FX rate delivered to the *Blue Chip* quarterly forecasts have been subject to topically oriented trend adjustment. In their earlier study Bofinger and Schmidt (2003) deliver similar results for the USD/EUR exchange rate.

Sign accuracy test Pierdzioch and Rülke (2015), for example, emphasize that despite possible biasedness of survey forecasts for exchange rates the predictions might still be useful in the sense of directional accuracy.[9] Hence, the directional accuracy of provided forecasts will be investigated using the widely acknowledged sign accuracy test (see Diebold and Lopez, 1996; Kolb and Stekler, 1996; Baghestani, 2008; Spiwoks, Bedke, and Hein, 2009; Baghestani, Arzaghi, and Kaya, 2015). There also exists a handful of studies focusing on the directional accuracy of exchange rate forecasts for Asian currencies. Tsuchiya and Suehara (2015) analyze the directional accuracy of USD/CNY forecasts also using *Consensus Economics* data and conclude that only 12M forecasts are useful. The authors investigate 1M, 3M, and 12M forecasts. Hence, the forecast horizons provided in the data set of this study might deliver additional insights regarding directional accuracy of USD/CNY forecasts. Pierdzioch and Rülke (2015) inter alia analyze FX forecasts for the Singapore dollar with horizons of one respectively three months. This paper follows the approach used for example by Spiwoks, Bedke, and Hein (2009). Table 3 below shows a two by two contingency table. The sum of the entries in N_{11} and N_{22} gives the number of correct forecasts for the directional change of the exchange rate. The number of incorrect forecasts is given by N_{21} and N_{12}, respectively.

Table 3: Sign accuracy test

Forecast \ Actual event	USD appreciates	USD depreciates	$\sum$
USD appreciates	N_{11}	N_{12}	$N_{1.}$
USD depreciates	N_{21}	N_{22}	$N_{2.}$
$\sum$	$N_{.1}$	$N_{.2}$	N

This table shows the two by two contingency table for the sign accuracy test (see Spiwoks, Bedke, and Hein, 2009).

It will be tested for significant differences regarding directional accuracy between a random directional prediction and the exchange rate forecast by performing a χ^2 test (see, for example, Diebold and Lopez, 1996; Nolte and Pohlmeier, 2007; Spiwoks, Bedke, and Hein, 2009; Baghestani, 2010; Tsuchiya and Suehara, 2015).

[9]The concept of rationality in the sense of unbiasedness and orthogonality will be presented in Section 4.2.

4.2. Rationality of foreign exchange forecasts

Finally, the rationality of predictions is an important attribute in the context of forecast evaluation. This is also true for exchange rate predictions. In the case of the USD/JPY exchange rate, for example, Frenkel, Rülke, and Stadtmann (2009) test for the rationality of FX survey forecasts by applying two criteria – namely, unbiasedness and orthogonality (i.e. efficiency). These criteria are common measures in the literature dealing with forecast evaluation (see also Ito, 1990; MacDonald and Marsh, 1996) and, hence, will be applied in this paper as well.

Test for unbiasedness Firstly, the test for unbiasedness will be used (see, for example, Hafer, Hein, and MacDonald, 1992, who apply the test for unbiasedness for futures market quotes, forward rates and survey forecasts for interest rates). Audretsch and Stadtmann (2005), Frenkel, Rülke, and Stadtmann (2009), Frenkel, Mauch, Rülke, et al. (2017), as well as Ince and Molodtsova (2017) investigate whether survey forecasts are unbiased predictors of future FX rates and come to different conclusions. Following, for example, Chinn and Frankel (1994), and more recently Ince and Molodtsova (2017) a simple linear regression model with the actual exchange rate change ($s_{t+h} - s_t$) as dependent variable and the expected exchange rate change ($\hat{s}_{t+h} - s_t$) as the independent variable will be estimated, where s_t is the log of the price of one US dollar in the foreign currency and the forecast of the exchanges at t for the horizon h is given as $\hat{s}_{t+h}$ (with the forecast horizons h = 3, 12, or 24 months).[10] The error term is given by u_{t+h}. For the test for unbiasedness the joint $H_0 : \alpha = 0$ and $\beta = 1$ has to be empirically verified respectively discarded (see, once again, Ince and Molodtsova, 2017).[11]

$$s_{t+h} - s_t = \alpha + \beta(\hat{s}_{t+h} - s_t) + u_{t+h} \tag{3}$$

Test for efficiency To test for orthogonality (see, for example, Frenkel, Rülke, and Stadtmann, 2009) the test for efficiency will be used. This test is also widely used in the context of forecast evaluation (see inter alia Simon, 1989; Leitner and Schmidt, 2006). Using this procedure, it is possible to verify whether the forecast errors are not related to information available at the forecast date (see Nordhaus, 1987; Frenkel, Rülke, and Stadtmann, 2009; Pancotto, Pericoli, and Pistagnesi, 2014; Frenkel, Mauch, Rülke, et al., 2017). Within the empirically context of this paper this information will be represented

[10] Due to the non-stationarity of the USD/CNY, USD/JPY, and USD/SGD exchange rates and corresponding forecasts regressions in levels would lead to distorted results (see Granger and Newbold, 1974; Mitchell and Pearce, 2007). Notwithstanding, due to the stationarity of the USD/HKD FX rate and the corresponding forecasts, to test for unbiasedness in the case of the USD/HKD forecasts, the test will be executed in levels.

[11] In addition to that null hypothesis the residuals of the linear model have to be tested for autocorrelation. The widely acknowledged Durbin Watson (DW) test is applied here (see Durbin and Watson, 1950, 1951)

by the last available monthly exchange rate change at the forecast date ($s_t - s_{\text{t-h}}$).

$$s_{t+h} - \hat{s}_{t+h} = \alpha + \beta(s_t - s_{t-h}) + u_{t+h} \tag{4}$$

Here, the joint null hypothesis is given by $H_0 : \alpha = 0$ and $\beta = 0$. Hence, if α and / or β are significantly different from zero there does not exist any empirical evidence for efficiency of forecasts (see, for example, Frenkel, Rülke, and Stadtmann, 2009; Ince and Molodtsova, 2017).

Cointegration tests In addition to the before mentioned tests for rationality this paper focuses on alternative measures of rationality. Cheung and Chinn (1998) propose to investigate the properties of the relevant time series. Following Cheung and Chinn (1998), as a necessary condition for rationality, the time series of the actual exchange rates and the corresponding forecast should share the same order of integration. In addition to that, there should exist long term relationships (i.e. cointegration) between the two time series (see, for example, Liu and Maddala, 1992; Cheung and Chinn, 1998). Two time series are said to be cointegrated when they share a common stochastic trend (see, for example, Granger, 1981; Engle and Granger, 1987; Lütkepohl and Krätzig, 2004). To test for cointegrating relationships the Johansen procedure will be used (see Johansen, 1991). In a further step it will be tested for Granger causal relationships. If there exists a cointegrating relationship between the two time series this can be seen as evidence for the existence of at least unidirectional Granger causality. In general, a time series x_t Granger causes time series y_t when past values of x_t provide additional content when forecasting y_t. (see, for example, Engle and Granger, 1987; Gelper and Croux, 2007); or phrased somewhat differently past values of x_t help to predict future values of y_t. Hence, for the FX forecasts to be useful predictors of future currency movements in this sense the survey forecasts ($\hat{s}_{t+h,t}$) should Granger cause the actual exchange rates (s_{t+h}).

Order of integration Given the initial results from section 4, the first necessary condition is fulfilled for the forecasts under investigation, since the PP test for unit roots indicates that the actual FX rate and the corresponding survey predictions share the same order of integration for all currency pairs (see Table 2 above). The USD/CNY, USD/JPY and USD/SGD FX rate and their corresponding forecasts are I(1) (i.e. non-stationary). These results correspond to earlier findings of Cheung and Chinn (1998), who have investigated 1M, 6M, and 12M forecasts for the USD/JPY, USD/CAD and the USD/DEM forecast. The USD/HKD FX rate and the corresponding forecasts are I(0) (i.e. stationary). Owing to the FX regime of a currency board, this finding is not surprising at all. [12]

[12]Due to the stationarity testing for cointegration makes no sense in the case of the USD/HKD FX rate. Hence, the cointegration tests are only excecuted for the USD/CNY, USD/JPY, and USD/SGD

5. Empirical evidence

Results Theil's U, DM test, TOTA coefficient The results of the forecast accuracy measures Theil's U, DM test as well as the TOTA coefficient are shown in Table 4.

Table 4: Measures of Forecast Accuracy

Exchange rate	Time series	Theil's U	pv DM test	TOTA
USD/CNY	3M forecast	0.78	0.11	0.99
	12M forecast	0.71	0.32	0.99
	24M forecast	0.86	0.69	0.93
USD/HKD	3M forecast	1.22	0.02	0.97
	12M forecast	1.03	0.81	0.86
	24M forecast	0.92	0.63	1.03
USD/JPY	3M forecast	1.13	0.01	0.85
	12M forecast	1.12	0.08	0.51
	24M forecast	0.86	0.20	0.10
USD/SGD	3M forecast	1.09	0.04	0.95
	12M forecast	1.05	0.58	0.81
	24M forecast	1.04	0.78	0.58

Reported are the results for the Theil's U measure, the Diebold Mariano (DM) test and the TOTA coefficient for the forecasts for USD/CNY, USD/HKD, USD/JPY, and USD/SGD with forecast horizons of 3, 12, and 24 months.

Based on the results of the Theil's U measure (see Table 4) the mean forecast for the USD/CNY FX rate[13] seems to be more accurate than the naïve prediction for all three forecast horizons. For the USD/HKD 24M, as well as the USD/JPY 24M forecast, the Theil's U measure indicates that the survey predictions are more accurate than the naïve forecast. For the remaining seven forecast time series a Theil's U measure larger than one implies that the naïve prediction is more accurate. Considering the results from the DM test procedure for equal forecast accuracy delivers rather sobering results: From the five forecast time series with a Theil's U < 1 the H_0 of equal forecast accuracy cannot be rejected on the 10% level. Even more than that, the H_0 of equal forecast accuracy has to be rejected for four forecast time series which have a Theil's U > 1 (i.e. USD/HKD

FX rates.

[13]The results for the USD/CNY FX rate presented in this Chapter for the period January 1999 to March 2017 are similar to the results for the time period August 2005 to March 2017 and, hence, will not be reported in order to preserve space. The results are available by request from the corresponding author.

3M, USD/JPY 3M, USD/JPY 12M and USD/SGD 3M). Hence, especially in the case of shorter forecast horizons, the naïve predictions seems to be the superior forecast. Interestingly, the superiority of the naïve predicition seems to be regime independent. As regards topically oriented trend adjustments, the only TOTA coefficient > 1 belongs to the USD/HKD 24M forecast. Especially in the case of the USD/JPY forecasts, the topically oriented forecasting behavior becomes obvious (i.e. $TOTA = 0.10$).

Results sign accuracy test The results of the sign accuracy test are presented in Table 5 below. On the 5%-level the USD/CNY as well as the USD/HKD forecasts are significantly better than a random forecast for all three forecast horizons. Hence, the results contradict the earlier findings of Tsuchiya and Suehara (2015). For the USD/JPY and USD/SGD predictions the results are more heterogeneous. For the 3M forecast for the USD/JPY as well as the USD/SGD no statistically significant difference from the random prediction regarding the sign accuracy has been found. The USD/JPY 12M forecast is worse than a random prediction for the direction of change, whereas the 12M forecast for USD/SGD outperforms the random prediction. The 24M forecasts both for the USD/JPY and for USD/SGD are better than a random prediction. As regards the regime dependent quality of FX forecasts it is noteworthy that the USD/CNY exchange rate as well as the USD/HKD exchange rate forecasts are consistently more sign accurate than the random prediction.

Table 5: Results sign accuracy test

Exchange rate	Time series	Test statistic	Test result
USD/CNY	3M forecast	18.15	+
	12M forecast	14.10	+
	24M forecast	4.00	+
USD/HKD	3M forecast	16.72	+
	12M forecast	42.32	+
	24M forecast	47.43	+
USD/JPY	3M forecast	1.42	*N/A*
	12M forecast	10.34	−
	24M forecast	3.46	+
USD/SGD	3M forecast	0.86	*N/A*
	12M forecast	5.04	+
	24M forecast	12.41	+

Reported are the test results for the sign accuracy test for USD/CNY, USD/HKD, USD/JPY and USD/SGD with the forecast horizons of 3, 12, and 24 months. The critical value on the 5%-level is 3.84. '+' indicates that the survey forecast is better than a random prediction. '−' indicates that the random prediction is better than the survey forecast. '*N/A*' indicates that no significant difference in sign accuracy exists.

Results test for unbiasedness The results of the test for unbiasedness are presented in Table 6 below. For the forecasts under investigation the joint $H_0 : \alpha = 0$ and $\beta = 1$ could not be empirically verified, because the requirement of nonexistence of autocorrelated residuals is not fulfilled. Hence, the predictions are biased for all forecasts respectively regimes under investigation. These results generally confirm earlier findings in the literature (see, for example, Frenkel, Rülke, and Stadtmann, 2009; Frenkel, Mauch, Rülke, et al., 2017; Ince and Molodtsova, 2017). Having said that, the empirical evidence does not allow any robust conclusion with regard to FX regime dependent differences of forecast rationality.

Table 6: Results unbiasedness test

Exchange rate	Time series	α	SE α	β	SE β	Test statistic	DW-test
USD/CNY	3M forecast	0.08	0.03	0.99	0.00	19.60	0.68
	12M forecast	0.73	0.07	0.90	0.01	120.78	0.13
	24M forecast	1.68	0.10	0.77	0.01	331.95	0.04
USD/HKD	3M forecast	0.38	0.14	0.81	0.07	46.20	0.37
	12M forecast	1.03	0.16	0.50	0.08	98.21	0.57
	24M forecast	1.05	0.12	0.49	0.06	173.24	0.48
USD/JPY	3M forecast	0.00	0.00	-0.14	0.14	68.32	0.68
	12M forecast	0.00	0.01	-0.13	0.14	66.54	0.19
	24M forecast	0.02	0.01	0.94	0.14	2.57	0.06
USD/SGD	3M forecast	0.00	0.00	0.09	0.13	48.56	0.68
	12M forecast	0.01	0.00	0.40	0.18	27.52	0.16
	24M forecast	0.03	0.01	1.05	0.29	35.32	0.07

Reported are the test results for the test for unbiasedness: The test statistic of the F-test as well as the DW-test, the coefficients α and β and the corresponding standard errors (SE) for USD/CNY, USD/HKD, USD/JPY and USD/SGD with the forecast horizons of 3, 12, and 24 months. The critical value on the 5%-level is 3.88.

Results test for efficiency The results of the efficiency tests are presented in Table 7 below. On the 5%-level empirical evidence for efficient forecasts has been found for the following forecasts: USD/CNY 3M, 12M and 24M, USD/HKD 12M, 24M, USD/SGD 12M, 24M. In contradiction to the results from the unbiasedness tests these findings are more convincing, because forecasts seem to be efficient for managed exchange rates.

Table 7: Results efficiency test

Exchange rate	Time series	Test statistic
USD/CNY	3M forecast	0.40
	12M forecast	0.03
	24M forecast	2.07
USD/HKD	3M forecast	12.2
	12M forecast	3.43
	24M forecast	0.85
USD/JPY	3M forecast	29.04
	12M forecast	14.39
	24M forecast	7.10
USD/SGD	3M forecast	9.92
	12M forecast	2.26
	24M forecast	2.70

Reported are the results of the efficiency tests' test statistic for USD/CNY, USD/HKD, USD/JPY and USD/SGD with the forecast horizons of 3, 12, and 24 months. The critical value on the 5%-level is 3.88.

Results cointegration tests Finally, as mentioned above, the procedure to test for the rationality of FX forecasts proposed by Cheung and Chinn (1998) will be applied. The results of the cointegration tests for the pairs of forecasts, and actual FX rate which are I(1) (i.e. the USD/JPY, the USD/SGD as well as the USD/CNY FX rate), are presented in the Tables 8 to 10 below. Interestingly, only the 24M forecast for the USD/JPY FX rate does not fulfill the condition of cointegration (see Table 10). For the remaining eight pairs of forecasts and actual FX rates empirical evidence for a cointegrating relationship has been found, i.e. the null hypothesis of no cointegration had to be rejected at least on the 5%-level.

Table 8: Results cointegration tests USD/CNY

	Test statistic	10%	5%	1%
3M forecast				
$r \leq 1$	0.91	6.50	8.18	11.65
$r = 0$	45.84	12.91	14.90	19.19
12M forecast				
$r \leq 1$	0.76	6.50	8.18	11.65
$r = 0$	25.62	12.91	14.90	19.19
24M forecast				
$r \leq 1$	0.77	6.50	8.18	11.65
$r = 0$	16.85	12.91	14.90	19.19

Reported are the test statistics and critical values for the cointegration tests for the USD/CNY forecast and the actual USD/CNY exchange rate for the 3M, 12M, and 24M forecast horizon. The H_0 of no cointegrating relationship is given by $r = 0$. The H_0 of at least one cointegrating relationship is given by $r \leq 1$.

Table 9: Results cointegration tests USD/JPY

	Test statistic	10%	5%	1%
3M forecast				
$r \leq 1$	3.53	6.50	8.18	11.65
$r = 0$	119.77	12.91	14.90	19.19
12M forecast				
$r \leq 1$	4.97	6.50	8.18	11.65
$r = 0$	22.50	12.91	14.90	19.19
24M forecast				
$r \leq 1$	3.44	6.50	8.18	11.65
$r = 0$	5.80	12.91	14.90	19.19

Reported are the test statistics and critical values for the cointegration tests for the USD/JPY forecast and the actual USD/JPY exchange rate for the 3M, 12M, and 24M forecast horizon. The H_0 of no cointegrating relationship is given by $r = 0$. The H_0 of at least one cointegrating relationship is given by $r \leq 1$.

Table 10: Results cointegration tests USD/SGD

	Test statistic	10%	5%	1%
3M forecast				
$r \leq 1$	0.96	6.50	8.18	11.65
$r = 0$	108.07	12.91	14.90	19.19
12M forecast				
$r \leq 1$	0.99	6.50	8.18	11.65
$r = 0$	23.53	12.91	14.90	19.19
24M forecast				
$r \leq 1$	2.00	6.50	8.18	11.65
$r = 0$	15.81	12.91	14.90	19.19

Reported are the test statistics and critical values for the cointegration tests for the USD/SGD forecast and the actual USD/SGD exchange rate for the 3M, 12M, and 24M forecast horizon. The H_0 of no cointegrating relationship is given by $r = 0$. The H_0 of at least one cointegrating relationship is given by $r \leq 1$.

Model estimation In a final evaluation, it will be tested for Granger causality using impulse response functions for the relevant bivariate vector autoregressive models (VAR) respectively bivariate vector error correction models (VECM). The results predetermine which model type is appropriate for the pairs of forecasts and actual FX rate. For the USD/CNY (3M, 12M, and 24M), the USD/JPY (3M, and 12M), as well as the USD/SGD (3M, 12M, and 24M) will be used. For the stationary USD/HKD time series (3M, 12M, and 24M) VAR models in levels will be used. And finally, for the USD/JPY 24M FX rates a VAR model in 1^{st} differences has to be applied due to the lack of a cointegrating relationship between the two non-stationary time series. Table 11 below summarizes the applied models. Lag lengths for the VAR models is determined using the Schwarz Criterion (SC) (see Schwarz et al., 1978). The vector error correction models have a lag of two.

Table 11: Estimated models for actual FX rates and corresponding forecasts

Exchange rate	Forecast horizon	Model	Lag
USD/CNY	3M	VECM	2
	12M	VECM	2
	24M	VECM	2
USD/HKD	3M	VAR in levels	3
	12M	VAR in levels	3
	24M	VAR in levels	3
USD/JPY	3M	VECM	2
	12M	VECM	2
	24M	VAR in 1^{st} differences	1
USD/SGD	3M	VECM	2
	12M	VECM	2
	24M	VECM	2

This table summarizes the estimated models (i.e. vector error correction model (VECM) as well as vector auto regressive (VAR) models) for the corresponding relationships between the forecasts and actual FX rates. Lag lengths are reported in the last column.

Useful financial market forecasts should deliver relevant information for actual future movements (see, for example, Schwarzbach, Kunze, Rudschuck, and Windels, 2012; Kunze, Wegener, Bizer, and Spiwoks, 2017). This should also be the case for exchange rates. As noted above the concept of Granger causality will be used to investigate the hypothesis whether forecasts under investigation are useful in a Granger sense.

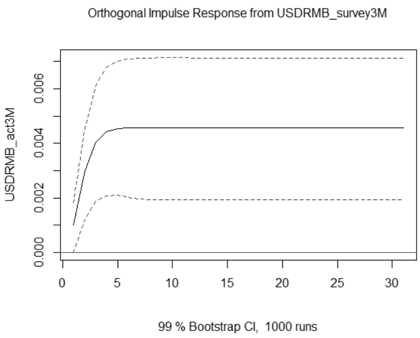
The empirical confirmation of a cointegrating relationship can be seen as evidence for the existence of at least unidirectional Granger causality. However, it is not clear whether the Granger causality is running from the forecast to the actual FX rate, from the actual FX rate to the forecast or if it is bidirectional. Furthermore, in the case of the VAR models no preliminary conclusions regarding Granger causal relationships are possible.

Impulse response analysis In an empirically setting impulse response functions are a useful tool to analyze these possible relationships between time series regarding Granger causality (see, for example, Basse and Reddemann, 2011). One advantage of impulse response functions stems from the fact that they are easy to interpret. The results of the impulse response analysis can be found in the Figures 1 to 4 below. The confidence intervals (with a confidence level of 99%) have been derived by bootstrapping (see Efron and Tibshirani, 1994) and 1,000 runs have been used. For the orthogonal impulses (see Sims, 1980) the length of the impulses is 30. Empirical evidence for a Granger causal relationship would be indicated by a significant response of the forecast time series to a shock to the actual times series or vice versa. A significant response is defined by a move of the response time series including the upper and lower confidence band above, or below the zero line (see Kunze, Wegener, Bizer, and Spiwoks, 2017).

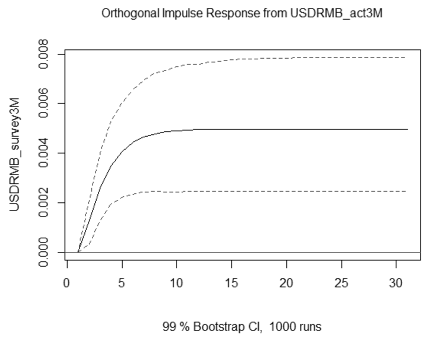
Following the IRF results for the USD/CNY 3M VECM shown in the panels (a) and (b) in Figure 1 it can be stated that there exists empirical evidence for bidirectional Granger causality between the time series of the FX forecast and the actual FX rate. Hence, short term forecasts for the USD/CNY FX rate are not only rational in the sense of the approach proposed by Cheung and Chinn (1998) (i.e. sharing the same order of integration), but it is also shown that the USD/CNY FX forecast with a three months horizon (3M) has forerunning properties regarding the USD/CNY exchange rate. However, for the longer term forecast horizons this finding is not sustainable (see panels c, d, e, and f in Figure 1). Instead, the 12M and 24M forecasts for the FX rates are only granger caused by the actual USD/CNY rate. This result may also be seen as supporting evidence for the results from the measures of forecast accuracy (i.e. especially the TOTA coefficient).

In the case of the VAR models for the forecast and actual time series for the USD/HKD exchange rates only one Granger causal relationship has been found (i.e. the response from the 3M USD/HKD forecast to an impulse from the the 3M USD/HKD actual rate as shown in panel b of Figure 2). For all three forecast horizons the USD/HKD FX rate predictions have no statistical significant forerunning properties for the actual FX rate. This may come as a surprise, because given the special properties of the currency board arrangement conventional wisdom would indicate that predicting the FX rate is easier. Having said that, it is important to note, that due to the stationarity of the time series under investigation for neither forecast horizon there exists a long term relationship between the actual rate and the forecast.

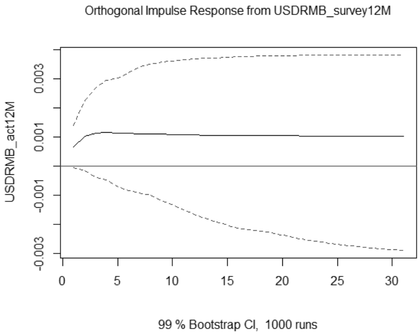
The interpretation of the results for the IRF for the USD/JPY as well as for the USD/SGD forecasts is rather straightforward. Based on the vector error correction models no evidence for forerunning properties of relevant forecast time series has been found (see Figure 3 panel a and c as well as Figure 4 panel a, c, and e). The same holds true for the VAR model in first differences for the USD/JPY exchange rate with a 24M forecast horizon (see Figure 3 panel e).



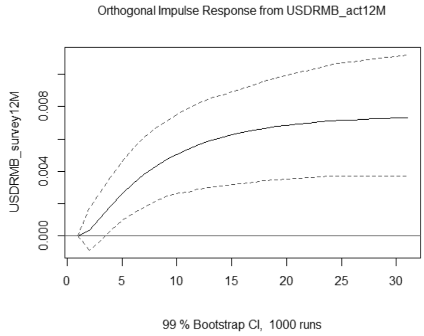
(a) 3M: Impulse by USD/CNY forecast



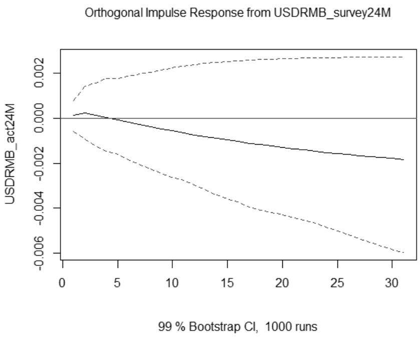
(b) 3M: Impulse by USD/CNY actual



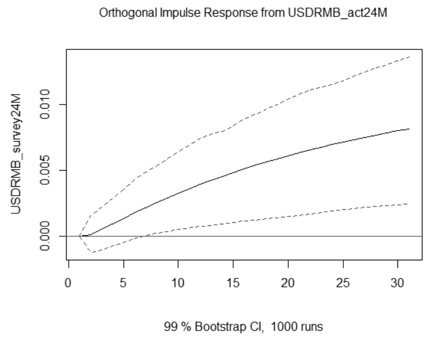
(c) 12M: Impulse by USD/CNY forecast



(d) 12M: Impulse by USD/CNY actual

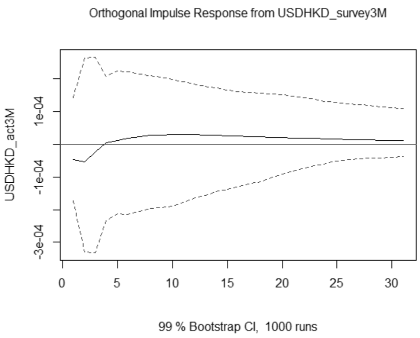


(e) 24M: Impulse by USD/CNY forecast

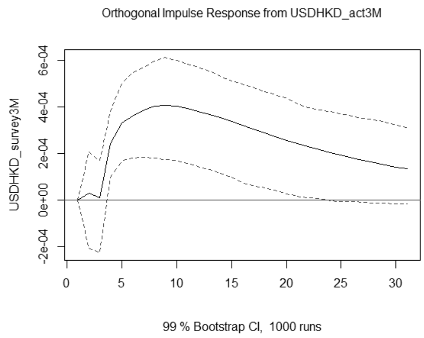


(f) 24M: Impulse by USD/CNY actual

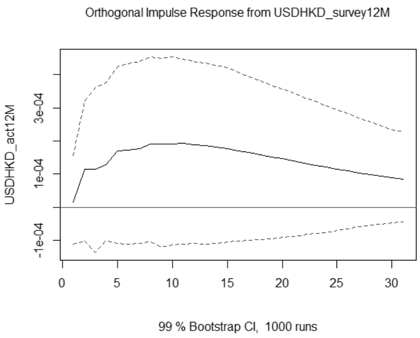
Fig. 1. The graphs show the impulse response functions for the USD/CNY forecast and the actual USD/CNY exchange rate for the 3M, 12M, and 24M forecast horizon.



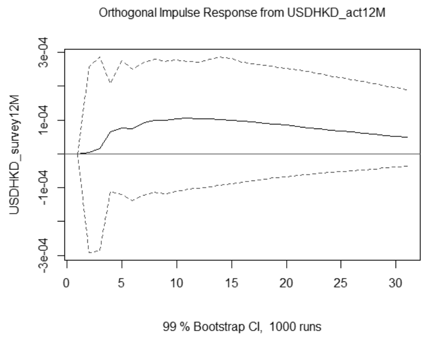
(a) 3M: Impulse by USD/HKD forecast



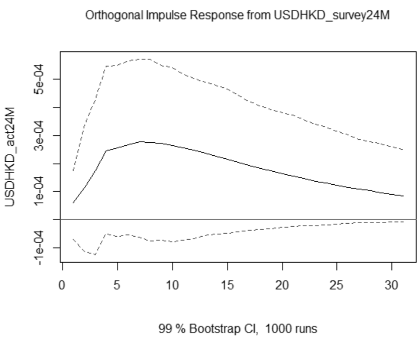
(b) 3M: Impulse by USD/HKD actual



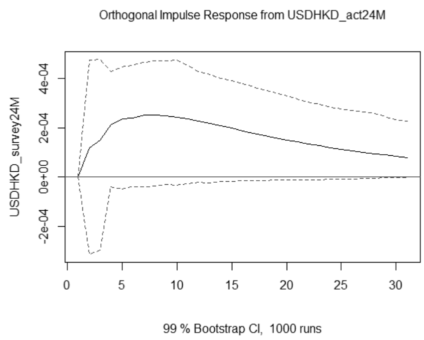
(c) 12M: Impulse by USD/HKD forecast



(d) 12M: Impulse by USD/HKD actual

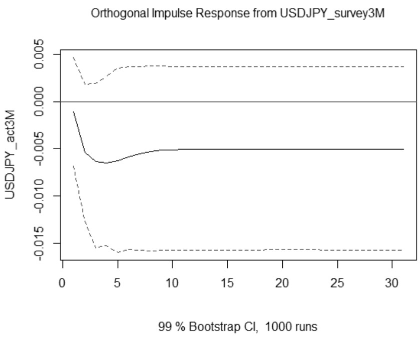


(e) 24M: Impulse by USD/HKD forecast

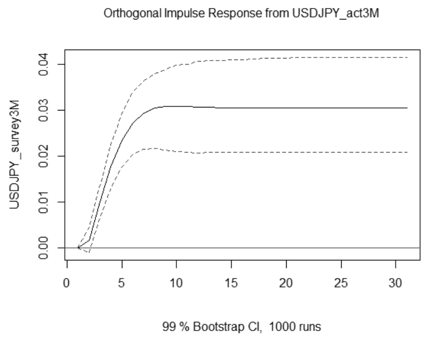


(f) 24M: Impulse by USD/HKD actual

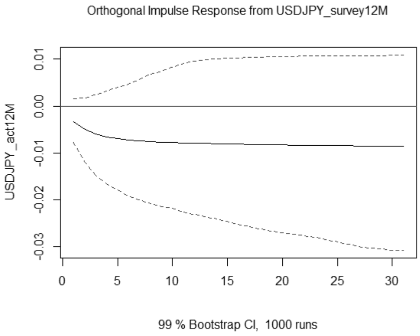
Fig. 2. The graphs show the impulse response functions for the USD/HKD forecast and the actual USD/HKD exchange rate for the 3M, 12M, and 24M forecast horizon.



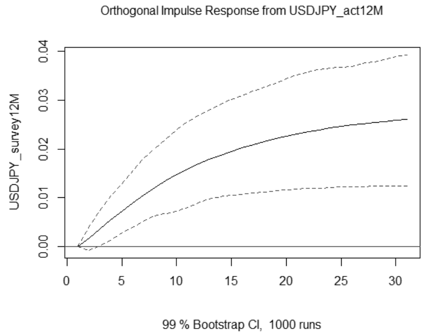
(a) 3M: Impulse by USD/JPY forecast



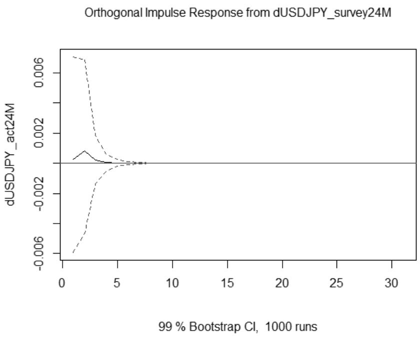
(b) 3M: Impulse by USD/JPY actual



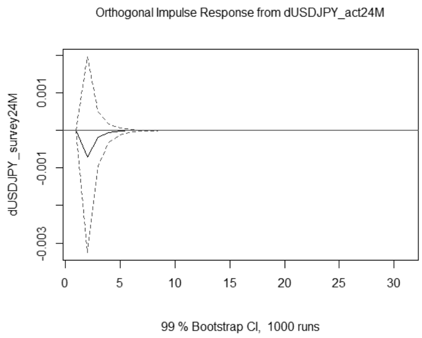
(c) 12M: Impulse by USD/JPY forecast



(d) 12M: Impulse by USD/JPY actual

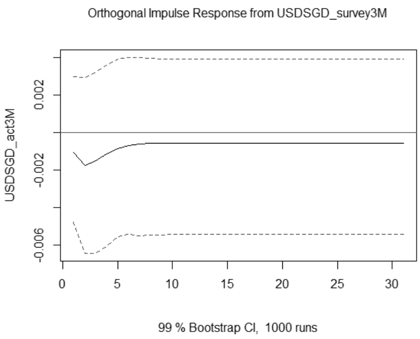


(e) 24M: Impulse by USD/JPY forecast

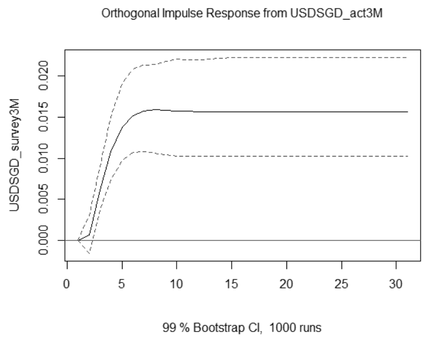


(f) 24M: Impulse by USD/JPY actual

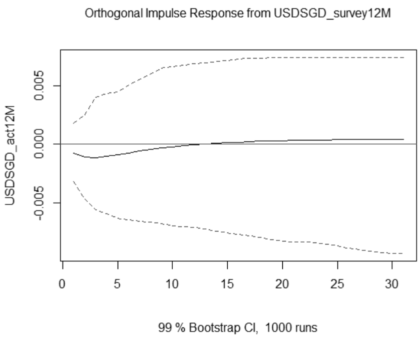
Fig. 3. The graphs show the impulse response functions for the USD/JPY forecast and the actual USD/JPY exchange rate for the 3M, 12M, and 24M forecast horizon.



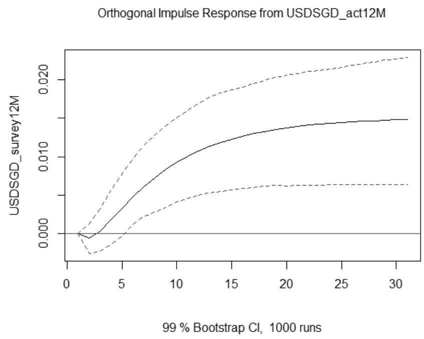
(a) 3M: Impulse by USD/SGD forecast



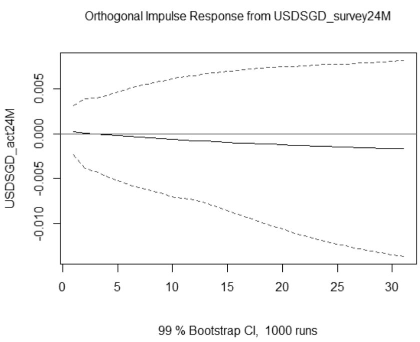
(b) 3M: Impulse by USD/SGD actual



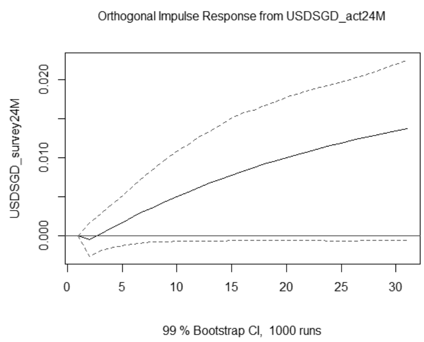
(c) 12M: Impulse by USD/SGD forecast



(d) 12M: Impulse by USD/SGD actual



(e) 24M: Impulse by USD/SGD forecast



(f) 24M: Impulse by USD/SGD actual

Fig. 4. The graphs show the impulse response functions for the USD/SGD forecast and the actual USD/SGD exchange rate for the 3M, 12M, and 24M forecast horizon.

6. Conclusions

In this paper monthly survey based exchange rate forecasts for the exchange rates of the Chinese yuan, the Hong Kong dollar, the Japanese yen as well as the Singapore dollar vis-à-vis the US dollar have been investigated regarding accuracy and rationality. Thereby, some interesting and relevant general results are worth mentioning. Rather strong empirical evidence for topically orientated (TOTA) forecasting behavior for all four exchange rates has been found. Furthermore, especially for shorter forecast horizons (i.e. 3M) the naïve *no change* prediction seems to outperform the survey forecasts with the three months forecasts for the USD/CNY exchange rate being the only exception. As regards sign accuracy for the forecasts for the USD/CNY and USD/HKD, empirical evidence supports the conclusion that the survey forecasts are more precise than a random prediction for all three forecast horizons under investigation. These results have been supported by the test for efficiency. Having said that, empirical evidence shows that all forecasts are irrational in the sense of the test for unbiasedness, due to the existence of autocorrelated residuals.

Applying an alternative framework to test for rationality proposed by Cheung and Chinn (1998), it has been shown that, for all four currency regimes the forecasts and actual exchange rates share the same order of integration. Furthermore, for the foreign exchange regimes with I(1) time series (i.e. USD/CNY, USD/JPY, and USD/SGD) empirical evidence supports the hypothesis of the existence of cointegrating relationships and, hence, a long term relation has been statistically validated. The only exception has been the USD/JPY forecast with a forecast horizon of 24 months. However, impulse response analysis for all currency regimes under investigation indicated that only the three months forecast for the managed USD/CNY exchange rate has forerunning properties. More generally, forerunning properties of the remaining survey forecasts under investigation are rather limited.

Furthermore, as regards regime dependent differences with respect to accuracy and rationality it has not been shown that exchange rates for fixed or closely managed exchange rates are in general easier to forecast. Despite the fact that no definite conclusion regarding FX regime dependent forecast accuracy respectively rationality of the survey predictions is suitable, the presented empirical evidence indicates that the forecasts for the managed Chinese exchange rate systems are closest to be being rational and forerunning, especially for the three months forecast horizons. Furthermore, also taking into account the results for the forecast evaluation of the USD/HKD exchange rate it can be stated that forecast accuracy – especially when it comes to sign accuracy – seems to be higher for stronger controlled exchange rate regimes. These findings are relevant both for the recipients of foreign exchange forecasts (e.g. corporate managers or politicians) and for the monetary respectively FX policy makers themselves.

After all, these findings may not be extraordinary surprising, as in managed foreign exchange regimes monetary policy authorities have to follow a self or government obliged framework. However, it has to be taken into account, that strongly regulated FX markets may be much more exposed to event risks and tail events, respectively, like for example currency crisis and, most importantly in the context of the results of this study, shifts in the FX regimes (see, for example, Husain, Mody, and Rogoff, 2005; Fiess and Shankar, 2009; Abildgren, 2014). Having said that, further research is necessary and should, particularly, focus on the predictability of regime shifts in managed exchange rate systems and currency boards, the credibility of fixed exchange rate regimes as well as the influence thereof on forecast accuracy and rationality, respectively.

References

Abildgren, K., 2014. Tail events in the FX markets since 1740. The Journal of Risk Finance 15, 294–311.

Aggarwal, R., 2013. Corporate management of foreign currency risk: Conceptual framework, policies, and strategies. International Finance: A Survey p. 447.

Aggarwal, R., Chen, X., Yur-Austin, J., 2011. Currency risk exposure of Chinese corporations. Research in International Business and Finance 25, 266–276.

Ahmed, S., Liu, X., Valente, G., 2016. Can currency-based risk factors help forecast exchange rates? International Journal of Forecasting 32, 75–97.

Andres, P., Spiwoks, M., 1999. Forecast quality matrix: A methodological survey of judging forecast quality of capital market forecasts. Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik) 219, 513–542.

Audretsch, D. B., Stadtmann, G., 2005. Biases in FX-forecasts: Evidence from panel data. Global Finance Journal 16, 99–111.

Avraham, D., Ungar, M., Zilberfarb, B.-Z., 1987. Are foreign exchange forecasts rational?: An empirical note. Economics Letters 24, 291–293.

Bacchetta, P., Mertens, E., Van Wincoop, E., 2009. Predictability in financial markets: What do survey expectations tell us? Journal of International Money and Finance 28, 406–426.

Baghestani, H., 2008. Federal Reserve versus private information: Who is the best unemployment rate predictor? Journal of Policy Modeling 30, 101–110.

Baghestani, H., 2010. Evaluating Blue Chip forecasts of the trade-weighted dollar exchange rate. Applied Financial Economics 20, 1879–1889.

Baghestani, H., Arzaghi, M., Kaya, I., 2015. On the accuracy of Blue Chip forecasts of interest rates and country risk premiums. Applied Economics 47, 113–122.

Basse, T., Reddemann, S., 2011. Inflation and the dividend policy of US firms. Managerial Finance 37, 34–46.

Beckmann, J., Czudaj, R., 2017a. Exchange rate expectations since the financial crisis: Performance evaluation and the role of monetary policy and safe haven. Journal of International Money and Finance 74, 283–300.

Beckmann, J., Czudaj, R., 2017b. The impact of uncertainty on professional exchange rate forecasts. Journal of International Money and Finance 73, 296–316.

Beine, M., Bénassy-Quéré, A., MacDonald, R., 2007. The impact of central bank intervention on exchange-rate forecast heterogeneity. Journal of the Japanese and International Economies 21, 38–63.

Bofinger, P., Schmidt, R., 2003. On the reliability of professional exchange rate forecasts. Financial Markets and Portfolio Management 17, 437–449.

Bordo, M. D., Choudhri, E. U., Fazio, G., MacDonald, R., 2017. The real exchange rate in the long run: Balassa-Samuelson effects reconsidered. Journal of International Money and Finance 75, 69–92.

Byrne, J. P., Korobilis, D., Ribeiro, P. J., 2016. Exchange rate predictability in a changing world. Journal of International Money and Finance 62, 1–24.

Calvo, G. A., Reinhart, C. M., 2002. Fear of floating. The Quarterly Journal of Economics 117, 379–408.

Cheung, Y.-W., Chinn, M. D., 1998. Integration, cointegration and the forecast consistency of structural exchange rate models. Journal of International Money and Finance 17, 813–830.

Cheung, Y.-W., Chinn, M. D., Fujii, E., 2007. The overvaluation of renminbi undervaluation. Journal of International Money and Finance 26, 762–785.

Cheung, Y.-W., Hui, C.-H., Tsang, A., 2016. The Renminbi central parity: An empirical investigation. CESifo Working Paper No. 5963 .

Cheung, Y.-W., Hui, C.-H., Tsang, A., 2017. The RMB central parity formation mechanism after august 2015: A statistical analysis. HKIMR Working Paper No. 06/2017 .

Chinn, M., Frankel, J., 1994. Patterns in exchange rate forecasts for twenty-five currencies. Journal of Money, Credit and Banking 26, 759–770.

Chit, M. M., Rizov, M., Willenbockel, D., 2010. Exchange rate volatility and exports: New empirical evidence from the emerging East Asian economies. The World Economy 33, 239–263.

Chow, H.-K., 2007. Singapore's exchange rate policy: Some implementation issues. The Singapore Economic Review 52, 445–458.

Chow, H. K., Lim, G., McNelis, P. D., 2014. Monetary regime choice in Singapore: Would a Taylor rule outperform exchange-rate management? Journal of Asian Economics 30, 63–81.

Cook, D., Yetman, J., 2014. Currency boards when interest rates are zero. Pacific Economic Review 19, 135–151.

De Grauwe, P., Markiewicz, A., 2013. Learning to forecast the exchange rate: Two competing approaches. Journal of International Money and Finance 32, 42–76.

Devereux, M. B., 2003. A tale of two currencies: The Asian crisis and the exchange rate regimes of Hong Kong and Singapore. Review of International Economics 11, 38–54.

Dick, C. D., MacDonald, R., Menkhoff, L., 2015. Exchange rate forecasts and expected fundamentals. Journal of International Money and Finance 53, 235–256.

Diebold, F. X., Lopez, J. A., 1996. 8 forecast evaluation and combination. Handbook of statistics 14, 241–268.

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

Dominguez, K. M., 1986. Are foreign exchange forecasts rational?: New evidence from survey data. Economics Letters 21, 277–281.

Dreger, C., Stadtmann, G., 2008. What drives heterogeneity in foreign exchange rate expectations: Insights from a new survey. International Journal of Finance & Economics 13, 360–367.

Duffy, G., Giddy, I., 1975. The random behavior of flexible exchange rates. Journal of International Business Studies 6, 1–32.

Dunis, C. L., Shannon, G., 2005. Emerging markets of south-east and central Asia: Do they still offer a diversification benefit? Journal of Asset Management 6, 168–190.

Durbin, J., Watson, G., 1950. Testing for serial correlation in least squares regression-i. Biometrika 37, 409–428.

Durbin, J., Watson, G. S., 1951. Testing for serial correlation in least squares regression. II. Biometrika 38, 159–177.

Efron, B., Tibshirani, R. J., 1994. An introduction to the bootstrap. CRC press.

Engle, R. F., Granger, C. W., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society pp. 251–276.

Fiess, N., Shankar, R., 2009. Determinants of exchange rate regime switching. Journal of International Money and Finance 28, 68–98.

Frankel, J. A., Rose, A. K., 1994. A survey of empirical research on nominal exchange rates. Tech. rep., National Bureau of Economic Research.

Fratzscher, M., 2009. What explains global exchange rate movements during the financial crisis? Journal of International Money and Finance 28, 1390–1407.

Frenkel, J. A., Mussa, M. L., 1980. The efficiency of foreign exchange markets and measures of turbulence. The American Economic Review 70, 374–381.

Frenkel, M., Mauch, M., Rülke, J.-C., et al., 2017. Forecaster rationality and expectation formation in foreign exchange markets: Do emerging markets differ from industrialized economies? Tech. rep., WHU-Otto Beisheim School of Management, Economics Group.

Frenkel, M., Rülke, J.-C., Stadtmann, G., 2009. Two currencies, one model? Evidence from the Wall Street Journal forecast poll. Journal of International Financial Markets, Institutions and Money 19, 588–596.

Gelper, S., Croux, C., 2007. Multivariate out-of-sample tests for Granger causality. Computational statistics & data analysis 51, 3319–3329.

Granger, C. W., 1981. Some properties of time series data and their use in econometric model specification. Journal of econometrics 16, 121–130.

Granger, C. W., Newbold, P., 1974. Spurious regressions in econometrics. Journal of econometrics 2, 111–120.

Hafer, R. W., Hein, S. E., MacDonald, S. S., 1992. Market and survey forecasts of the three-month treasury-bill rate. Journal of Business pp. 123–138.

Hamada, K., Hayashi, F., 1985. Monetary policy in postwar Japan. Monetary Policy in Our Times pp. 83–121.

Hayakawa, K., Kimura, F., 2009. The effect of exchange rate volatility on international trade in East Asia. Journal of the Japanese and International Economies 23, 395–406.

Heiden, S., Klein, C., Zwergel, B., 2013. Beyond fundamentals: Investor sentiment and exchange rate forecasting. European Financial Management 19, 558–578.

Hernandez, L., Montiel, P. J., 2003. Post-crisis exchange rate policy in five Asian countries: Filling in the "hollow middle"? Journal of the Japanese and International Economies 17, 336–369.

Herwartz, H., Schlüter, S., 2017. On the predictive information of futures' prices: A wavelet-based assessment. Journal of Forecasting 36, 345–356.

Ho, C., 2002. A survey of the institutional and operational aspects of modern-day currency boards. BIS Working Paper 110.

Husain, A. M., Mody, A., Rogoff, K. S., 2005. Exchange rate regime durability and performance in developing versus advanced economies. Journal of monetary economics 52, 35–64.

Hutchison, M., Walsh, C. E., 1992. Empirical evidence on the insulation properties of fixed and flexible exchange rates: The Japanese experience. Journal of International Economics 32, 241–263.

Hyndman, R. J., Koehler, A. B., 2006. Another look at measures of forecast accuracy. International journal of forecasting 22, 679–688.

Ince, O., Molodtsova, T., 2017. Rationality and forecasting accuracy of exchange rate expectations: Evidence from survey-based forecasts. Journal of International Financial Markets, Institutions and Money 47, 131–151.

Ito, T., 1990. Foreign exchange rate expectations: Micro survey data. The American Economic Review pp. 434–449.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica: Journal of the Econometric Society pp. 1551–1580.

Jongen, R., Verschoor, W. F., Wolff, C. C., Zwinkels, R. C., 2012. Explaining dispersion in foreign exchange expectations: A heterogeneous agent approach. Journal of Economic Dynamics and Control 36, 719–735.

Kan, Y. Y., 2017. Why RMB should be more flexible. Journal of Financial Economic Policy 9, 156–173.

Kilian, L., Taylor, M. P., 2003. Why is it so difficult to beat the random walk forecast of exchange rates? Journal of International Economics 60, 85–107.

Klein, M. W., Shambaugh, J. C., 2008. The dynamics of exchange rate regimes: Fixes, floats, and flips. Journal of International Economics 75, 70–92.

Kohlscheen, E., 2014. The impact of monetary policy on the exchange rate: A high frequency exchange rate puzzle in emerging economies. Journal of International Money and Finance 44, 69–96.

Kokenyne, A., Veyrune, R., Habermeier, M. K. F., Anderson, H., 2009. Revised system for the classification of exchange rate arrangements. No. 9-211, International Monetary Fund.

Kolb, R., Stekler, H. O., 1996. How well do analysts forecast interest rates? Journal of Forecasting 15, 385–394.

Kunze, F., Wegener, C., Bizer, K., Spiwoks, M., 2017. Forecasting European interest rates in times of financial crisis – What insights do we get from international survey forecasts? Journal of International Financial Markets, Institutions and Money 48, 192–205.

Leitch, G., Tanner, J. E., 1991. Economic forecast evaluation: Profits versus the conventional error measures. The American Economic Review pp. 580–590.

Leitner, J., Schmidt, R., 2006. A systematic comparison of professional exchange rate forecasts with the judgemental forecasts of novices. Central European Journal of Operations Research 14, 87–102.

Liu, P. C., Maddala, G. S., 1992. Rationality of survey data and tests for market efficiency in the foreign exchange markets. Journal of International Money and Finance 11, 366–381.

Lütkepohl, H., Krätzig, M., 2004. Applied time series econometrics. Cambridge university press.

MacDonald, R., Marsh, I. W., 1996. Currency forecasters are heterogeneous: Confirmation and consequences. Journal of International Money and Finance 15, 665–685.

MacDonald, R., Nagayasu, J., 2015. Currency forecast errors and carry trades at times of low interest rates: Evidence from survey data on the yen/dollar exchange rate. Journal of International Money and Finance 53, 1–19.

Meese, R. A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: Do they fit out of sample? Journal of international economics 14, 3–24.

Mitchell, K., Pearce, D. K., 2007. Professional forecasts of interest rates and exchange rates: Evidence from the Wall Street Journal's panel of economists. Journal of Macroeconomics 29, 840–854.

Moosa, I., Li, L., 2017. The mystery of the Chinese exchange rate regime: Basket or no basket? Applied Economics 49, 349–360.

Nolte, I., Pohlmeier, W., 2007. Using forecasts of forecasters to forecast. International Journal of Forecasting 23, 15–28.

Nordhaus, W. D., 1987. Forecasting efficiency: Concepts and applications. The Review of Economics and Statistics pp. 667–674.

Obstfeld, M., Rogoff, K., 1995. The mirage of fixed exchange rates. Journal of Economic Perspectives 9, 73–96.

Pancotto, F., Pericoli, F. M., Pistagnesi, M., 2014. Overreaction in survey exchange rate forecasts. Journal of Forecasting 33, 243–258.

Patnaik, I., Shah, A., Sethy, A., Balasubramaniam, V., 2011. The exchange rate regime in Asia: From crisis to crisis. International Review of Economics & Finance 20, 32–43.

Phillips, P. C., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.

Pierdzioch, C., Rülke, J.-C., 2015. On the directional accuracy of forecasts of emerging market exchange rates. International Review of Economics & Finance 38, 369–376.

Reitz, S., Stadtmann, G., Taylor, M. P., 2010. The effects of Japanese interventions on FX-forecast heterogeneity. Economics Letters 108, 62–64.

Ryan, L., Whiting, B., 2017. Multi-model forecasts of the West Texas Intermediate crude oil spot price. Journal of Forecasting 36, 395–406.

Rülke, J. C., Frenkel, M. R., Stadtmann, G., 2010. Expectations on the yen/dollar exchange rate–Evidence from the Wall Street Journal forecast poll. Journal of the Japanese and International Economies 24, 355–368.

Rülke, J.-C., Pierdzioch, C., 2013. Currency crises, uncertain fundamentals and private-sector forecasts. Applied Economics Letters 20, 489–494.

Schwarz, G., et al., 1978. Estimating the dimension of a model. The annals of statistics 6, 461–464.

Schwarzbach, C., Kunze, F., Rudschuck, N., Windels, T., 2012. Asset management in the German insurance industry: The quality of interest rate forecasts. Zeitschrift für die gesamte Versicherungswissenschaft 101, 693–703.

Simon, D. P., 1989. The rationality of federal funds rate expectations: evidence from a survey. Journal of Money, Credit and Banking 21, 388–393.

Sims, C. A., 1980. Macroeconomics and reality. Econometrica: Journal of the Econometric Society pp. 1–48.

Siregar, R. Y., Har, C. L., et al., 2001. Economic fundamentals and managed floating exchange rate regime in Singapore. Journal of Economic Development 26, 133–148.

Sosvilla-Rivero, S., Ramos-Herrera, M. d. C., 2013. On the forecast accuracy and consistency of exchange rate expectations: The Spanish PwC survey. Applied Economics Letters 20, 107–110.

Spiwoks, M., Bedke, N., Hein, O., 2009. The pessimism of Swiss bond market analysts and the limits of the sign accuracy test–an empirical investigation of their forecasting success between 1998 and 2007. International Bulletin of Business Administration 4, 6–19.

Spiwoks, M., Hein, O., 2007. Die Währungs-, Anleihen-und aktienmarktprognosen des Zentrums für Europäische Wirtschaftsforschung. AStA Wirtschafts-und Sozialstatistisches Archiv 1, 43–52.

Takagi, S., 2007. Managing flexibility: Japanese exchange rate policy, 1971–2007. The Singapore Economic Review 52, 335–361.

Ter Ellen, S., Verschoor, W. F., Zwinkels, R. C., 2013. Dynamic expectation formation in the foreign exchange market. Journal of International Money and Finance 37, 75–97.

Theil, H., 1955. Who forecasts best? International Economic Papers 5, 194–199.

Thorbecke, W., 2008. The effect of exchange rate volatility on fragmentation in East Asia: Evidence from the electronics industry. Journal of the Japanese and International Economies 22, 535–544.

Tian, L., Chen, L., 2013. A reinvestigation of the new RMB exchange rate regime. China Economic Review 24, 16–25.

Tse, Y. K., Yip, P. S., 2006. Exchange-rate systems and interest-rate behaviour: The experience of Hong Kong and Singapore. International Review of Economics & Finance 15, 212–227.

Tsuchiya, Y., Suehara, S., 2015. Directional accuracy tests of Chinese renminbi forecasts. Journal of Chinese Economic and Business Studies 13, 397–406.

Von Spreckelsen, C., Kunze, F., Windels, T., von Mettenheim, H.-J., 2014. Forecasting renminbi quotes in the revised Chinese FX market – can we get implications for the onshore/offshore spread-behaviour? International Journal of Economic Policy in Emerging Economies 7, 66–76.

Woo, J., 2016. Business and politics in Asia's key financial centres. Springer.

Appendix

Table 12: China FX arrangement based on IMF's Annual Report on Exchange Arrangements and Exchange Restrictions

Arrangement	AREAER reports	Important announcements / changes
Conventional pegged arrangement	2000 - 2005	○ July 21st, 2005: People's Bank of China (PBOC) revalued USD/CNY to 8.11; CNY FX rate will be determined by an undisclosed basket of currencies
Crawling peg	2006 - 2007	○ effective August 1st. 2006 IMF classifies FX arrangement as crawling peg
Stabilized arrangement	2008 - 2009	○ from April 30st, 2008 IMF FX arrangement was classified as crawl-like arrangement due to changes in IMF classification
Crawl-like arrangement	2010 - 2015	○ effective June 1st, 2008 IMF classifies FX arrangement as stabilized arrangement
Other managed arrangement	2016	○ effective June 21st, 2010 the de facto exchange rate was reclassified to a crawl-like arrangement
		○ April 16th, 2012: USD/CNY trading band was officially widened from +/-0.5% to 1.0%
		○ March 17th, 2014: USD/CNY trading band was officially widened from +/-1.0% to 2.0%
		○ effective December 24th, 2014 the de facto exchange rate was reclassified to other managed arrangement

IMF's Annual Report on Exchange Rate Arrangements and Exchange Restrictions 2000-2016

Table 13: Japan FX arrangement based on IMF's Annual Report on Exchange Arrangements and Exchange Restrictions

Arrangement	AREAER reports	Important announcements / changes
Independently floating	2000 - 2007	∘ from April 30st, 2008 IMF FX arrangement was classified as free floating due to changes in IMF classification
Free floating	2008 - 2016	∘ September 15th, 2010: Ministry of Finance intervened in FX market
		- March 18th, August 4th, October 31st, 2011 Ministry of Finance intervened in FX market

IMF's Annual Report on Exchange Rate Arrangements and Exchange Restrictions 2000-2016

Table 14: Hong Kong FX arrangement based on IMF's Annual Report on Exchange Arrangements and Exchange Restrictions

Arrangement	AREAER reports	Important announcements / changes
Currency board	2000 - 2016	∘ May 18th, 2005: Hong Kong Monetary Authority (HKMA) established a trading band of USD/HKD 7.75 to USD/HKD 7.85

IMF's Annual Report on Exchange Rate Arrangements and Exchange Restrictions 2000-2016

Table 15: Singapore FX arrangement based on IMF's Annual Report on Exchange Arrangements and Exchange Restrictions

Arrangement	AREAER reports	Important announcements / changes
Managed floating with no pre-announced path for the exchange rate	2000 - 2005	∘ effective January 1, 2006, the de facto exchange rate arrangement has been reclassified retroactively to other managed arrangement
Other managed arrangement	2006 - 2009	∘ effective April 14th, 2010 arrangement was (retroactively) reclassified to crawl-like arrangement
Crawl-like arrangement	2010 - 2013	∘ effective September 12th, 2011 arrangement was reclassified to other managed arrangement
Stabilized arrangement	2014 - 2016	∘ effective November 9th, 2011 arrangement was reclassified to crawl-like arrangement
		∘ effective January 1st, 2013 arrangement was reclassified to stabilized arrangement

IMF's Annual Report on Exchange Rate Arrangements and Exchange Restrictions 2000-2016