Booms and Busts in Asset Prices: Risk Modeling, Bubble Detection, and the Role of Monetary Policy

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
doctor rerum politicarum
(Dr. rer. pol.)

am Fachbereich Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Benjamin Beckers

Berlin, 2017

Erstgutachter:	Prof. Dr. Helmut Lütkepohl *Bundesbankprofessur für Methoden* *der empirischen Wirtschaftsforschung* *Freie Universität Berlin*
Zweitgutachter:	Prof. Dr. Kerstin Bernoth *Professor of Economics* *Hertie School of Governance und DIW Berlin*

Tag der Disputation: 30. Juni 2017

Acknowledgments

I am deeply grateful for the immense support of many people without whom the completion of this dissertation would not have been possible.

First and foremost, I thank my supervisor Helmut Lütkepohl for his continuous guidance and immeasurable kindness over the past years. I am grateful for his critical feedback whenever I needed his advice, and his unmatched dedication to sharing his immense wealth of knowledge. I am also deeply indebted to Kerstin Bernoth for her valuable guidance, especially during the early stages of this project. Furthermore, I enjoyed our fruitful collaboration on the forth chapter of this thesis. Two other chapters of this dissertation also benefited from the work and input of my co-authors. I thus want to express my deepest gratitude to Helmut Herwartz, Moritz Seidel, and Dirk Ulbricht.

I am very thankful for the opportunity to complete this thesis whilst working as a researcher at the Department of Macroeconomics of DIW Berlin, and as a member of the DIW Graduate Center. I have benefited greatly from the stimulating research environment at DIW, which would not have been the same without the presence of so many wonderful colleagues. I want to especially thank Michael Hachula, Christoph Große Steffen, Maximilian Podstawski, Pablo Anaya, Daniel Bierbaumer, Ferdinand Fichtner, Konstantin Kholodilin, Simon Junker, Philipp König, Claus Michelsen, Michele Piffer, Malte Rieth, Annika Schnücker, and the entire GC team. I also want to thank Dieter Nautz, Mathias Trabandt, Lars Winkelmann and the entire Chairs of Econometrics and Macroeconomics at Freie Universität Berlin for providing me with valuable feedback and inspiration at their seminars and workshops. Moreover, I thank Samya Beidas-Strom as well as Knut Are Aastveit and André Anundsen for inviting me to visit the International Monetary Fund and Norges Bank. Our joint work and both visits have provided me with invaluable experiences and inspirations for my future career. My work was in large parts funded by the German Research Foundation (DFG), which I gratefully acknowledge.

Finally, I thank my family for their enduring support and encouragement during both heights and lows. Without them, I would not have had the possibility to even commence this journey. Likewise, I am also deeply grateful for the support of my fellow students and my friends outside of economics. I enjoyed deepening our (economic) discussions at many social activities over the occasional beer(s), but most importantly, I thank you for keeping me sane and allowing me to forget about economic research at times. Most of all, I want to thank Annika for her unconditional support and loving encouragement over the past three years. I am excited about our next joint adventures to begin.

Berlin, May 2017 *Benjamin Beckers*

Erklärung zu Ko-Autorenschaften

Diese Dissertation besteht aus vier (Arbeits-)Papieren, von denen eines in Zusammenarbeit mit zwei Koautoren und zwei in Zusammenarbeit mit jeweils einer Koautorin/einem Koautoren entstanden sind. Der Eigenanteil an Konzeption, Durchführung und Berichtsabfassung der Kapitel lässt sich folgendermaßen zusammenfassen:

- ***"Risk Forecasting in (T)GARCH Models with Uncorrelated Dependent Innovations"*** mit Helmut Herwartz und Moritz Seidel

  Eigenanteil: 33 Prozent

  Dieses Papier ist 2017 in der Fachzeitschrift *Quantitative Finance*, Vol. 17(1), 121–137, `http://dx.doi.org/10.1080/14697688.2016.1184303`, erschienen.

- ***"Detecting Asset Price Bubbles in Real-Time through Indicator Combinations"***

  Eigenanteil: 100 Prozent

  Eine ältere Version dieses Kapitels ist als DIW Berlin Discussion Paper 1496/2015 erschienen.

- ***"Predicting Output with Real-Time Bubble Indicators*** mit Dirk Ulbricht

  Eigenanteil: 50 Prozent

- ***"Monetary Policy and Mispricing in Stock Markets"*** mit Kerstin Bernoth

  Eigenanteil: 50 Prozent

  Eine ältere Version dieses Kapitels ist als DIW Berlin Discussion Paper 1605/2016 erschienen.

Erklärung zu Vorarbeiten

Ich weise darauf hin, dass sich meine Masterarbeit mit dem Titel "The expected shortfall of speculative assets – modelling approaches by means of copula distributions" ebenfalls mit der Beschreibung von Abhängigkeiten in (T)GARCH Modellinnovationen mit Hilfe von Copulaverteilungen, sowie der Verbesserung von Risikoprognosen beschäftigt. Das erste Kapitel dieser Dissertation unterscheidet sich jedoch deutlich von dieser Arbeit und geht an mehreren Punkten weit über sie hinaus. Unter anderem überprüfen wir, ob die Modellinnovationen frei von serieller Korrelation und verbliebener Heteroskedastizität sind. Anschließend erweitern wir das Modellierungsspektrum erheblich in dem wir Student-t-Verteilungen sowohl in der TGARCH-Schätzung als auch für die Randverteilungen der Copulas berücksichtigen. In der Modellierung der bedingten Verteilungen von Modellinnovationen mit Copulaverteilungen stellen wir nun auch sicher, dass diese einer skalierten Martingaldifferenz folgen was einen weiteren Standardisierungsschritt im Simulationsprozess erfordert. Schließlich vergleichen wir die Risikoprognosen der vorgeschlagenen Copulamodelle auch mit einem nichtparametrischen Schätzansatz sowie mit Modellkombinationen, und evaluieren den Prognosegehalt aller Modelle ebenfalls für Value-at-Risk Prognosen.

Contents

List of Figures

List of Tables

List of Abbreviations

ADF Augmented Dickey-Fuller (test)
AIC Akaike Information Criterion
AR Autoregressive (model)
ARX Autoregressive model with exogenous variable(s)
ARCH Autoregressive Conditional Heteroskedasticity
BDS Brock-Dechert-Scheinkman (test)
BIC Bayesian Information Criterion
BPS Basis Points
BSADF Backward Supremum Augmented Dickey-Fuller (test)
cCDF conditional Cumulative Densitiy Function
CDF Cumulative Densitiy Function
CPI Consumer Price Index
cVaR conditional Value-at-Risk
C-VAR Constant coefficient Vector Autoregression
cES conditional Expected Shortfall
DDM Dividend Discount Model
DFM Dynamic Factor Model
DFM ELA Dynamic Factor Model with elastic net preselection
DM Diebold-Mariano (test)
DQ Dynamic Quantile (test)
ES Expected Shortfall
FED Federal Reserve Bank
FV Fundamental Value
FWE Familywise Error Rate
GARCH Generalized Autoregressive Conditional Heteroskedasticity (model)
GDP Gross Domestic Product

GFC Global Financial Crisis
GIRF Generalized Impulse Response Function
HP Hodrick-Prescott
IID Identically and Independently Distributed
IP Industrial Production
IPT Total Industrial Production Index
IRF Impulse Response Function
KPSS Kwiatkowski-Phillips-Schmidt-Shin (test)
LASSO Least Absolute Shrinkage and Selection Operator
LATW Leaning Against The Wind
LB Ljung-Box statistic
LM Lagrange Multiplier (test)
LR Likelihood Ratio (test)
MAE Mean Absolute Error
MC Monte Carlo
MCS Model Confidence Set
MDS Martingale Difference Sequence
ML Maximum Likelihood
NBER National Bureau for Economic Research
NW Nadaraya Watson estimator
PP Phillips-Perron (test)
PtD Price-to-Dividend
PtI Price-to-Income
QML Quasi Maximum Likelihood
RL Regulator's Loss
RMSPE Root Mean Square Prediction Error
sMDS scaled Martingale Difference Sequence
SVAR Structural Vector Autoregression
TGARCH Threshold Generalized Autoregressive Conditional Heteroskedasticity (model)
TVC Time-Varying Coefficients
U.S. United States (of America)
VaR Value-at-Risk
VAR Vector Autoregression
ZLB Zero Lower Bound

Introduction and Overview

Economic history has witnessed numerous excessive booms and busts in asset prices. Frequently, the burst of such speculative asset price bubbles has challenged not only the stability of the financial system, but has also been followed by severe economic contractions. Therefore, boom and bust cycles in asset prices have long attracted the attention of financial market participants, economists, and policymakers alike. The past two decades have brought about two prime examples for such cycles: the dot-com stock market bubble in the late 1990's and the U.S. housing market bubble in the 2000's. The latter in particular has provided a forceful reminder about the risks to financial stability and the adverse real economic consequences that may arise from the sudden burst of an asset price bubble. Following the housing market's peak in early 2006, falling house prices and a sharp increase in mortgage default rates triggered the global financial crisis (GFC). By the end of 2007, this crisis spilled over to the real economy which entered the longest and deepest recession since the Great Depression of 1929.

Consequently, the GFC has contested the understanding of macro- and financial economists of the risks inherent in asset markets, and of the linkages between the financial system and the real economy. As a result, the crisis has sparked intense debates about pre-crisis economic and financial policies, in particular with regard to financial market regulation and the role of monetary policy in amplifying or dampening asset price cycles. This dissertation consists of four chapters that empirically address some of the challenges emphasized by the recent booms and busts in asset prices, and thereby contributes to some of the initiated policy debates.

Specifically, the GFC has uncovered the failure of investors' internal risk models to accurately capture the implied tail risks of serially dependent, negative asset returns. The resulting underestimation of market risks has contributed to too low capital

buffers held by investors. Following the crisis, the Basel Committee on Banking Supervision has hence initiated a review of the internal risk models approach aiming to ensure a more prudent capture of tail risk and capital adequacy in financial stress situations (Basel Committee on Banking Supervision, 2009, 2016). Against this background, Chapter 1 offers a new approach for improving models commonly employed to describe and predict the evolution of financial market volatility and market risk. Thereby, our approach allows to better capture tail risks and serial dependency in speculative asset returns.

Furthermore, the GFC has highlighted the adverse real economic consequences arising from a bursting asset price bubble. As a result, this has stimulated debates on how to better contain emerging asset price bubbles and thereby promote financial and economic stability in the future. A particular set of policies that has since regained popularity are countercyclical macroprudential policies. Moreover, the pre-crisis conduct of an asymmetric, "benign neglect" monetary policy approach to emerging asset price bubbles is strongly debated. This approach builds on the notion that a central bank cannot and should not attempt to constrain the boom, but instead should only aim at promoting financial stability after the bust by cutting interest rates. Challenging this view, proponents of an active role of monetary policy have called on central banks to instead "lean against the wind" of asset price bubbles by preemptively raising interest rates early after their emergence.

For the appropriate design and use of countercyclical financial market policies in general, and a leaning against the wind monetary policy in particular, several important conditions must be fulfilled, though. First, policymakers must be able to correctly assess that an observed asset price appreciation is indeed excessive and warrants policy action. This prerequisite of detecting asset price bubbles has long been deemed impossible, especially when to be performed in real-time (Stiglitz, 1990; Kohn, 2006). In contrast to this view, Chapter 2 argues that recently developed econometric tests by Phillips et al. (2011, 2015) can be fruitfully combined to provide accurate real-time warning indicators about emerging asset price bubbles. To warrant a monetary policy response, however, these bubble signals must, secondly, contain predictive information for the central bank's real economic target variables. Here, Chapter 3 shows that the most accurate real-time bubble indicators developed in Chapter 2 also provide valuable information for forecasting future output

growth. Finally, in order to be able to lean against an emerging asset price bubble, it is required that central banks can lower asset mispricing by raising interest rates. Chapter 4 provides evidence that a monetary policy tightening indeed lowers stock prices beyond what the policy-induced changes of their fundamental values imply.

The four chapters of this dissertation are based on four individual papers. In the following, I briefly summarize their main results and their contributions to the literature.

Chapter 1: *Risk Forecasting in (T)GARCH Models with Uncorrelated Dependent Innovations*

The first chapter is based on a joint paper with Helmut Herwartz and Moritz Seidel. In this study we explore whether model residuals from the class of (threshold) generalized autoregressive conditional heteroskedasticity ((T)GARCH) models are characterized by serial dependence, which could potentially be used to enhance conventional risk forecasts. We find that these residuals are hardly independent and identically distributed (IID), but instead show forms of higher order serial dependence. This suggests that TGARCH models commonly employed for predicting market risk of speculative asset returns do not use all available information for their forecast. We propose two strategies to quantify the serial dependence structures between model innovations, a nonparametric estimation approach and a flexible modeling approach based on standardized copula distributions. We show that these strategies more accurately describe the in-sample dependence patterns between consecutive innovations, and outperform conventional TGARCH model predictions for the conditional Value-at-Risk (cVaR) and the conditional Expected Shortfall (cES) at the relevant risk levels outlined by the Basel Committee on Banking Supervision (2016).

Methodologically, we contribute to the literature on modeling univariate heteroskedastic time series by augmenting the commonly employed (T)GARCH model class developed by Engle (1982), Bollerslev (1986), and Glosten et al. (1993) to account for higher order serial dependence in the models' residuals. Since their introduction, these models have been extended along various dimensions, for example to describe other forms of non-linear behavior in the volatility process, or to allow for leptokurtic, non-normal conditional distributions of (multivariate) model residuals (e.g. Bollerslev, 1990). However, the assumption that the model innovations

are IID has rarely been challenged. For instance, until now it has only been suggested to implement conditional skewness (e.g. Harvey and Siddique, 1999) and kurtosis (Brooks et al., 2005) in model innovations by conditioning their distribution on lagged returns and the estimated variance. Yet, none of the extensions prior to our study has considered the standardized residuals to exhibit an explicit serial dependence structure, i.e. a dependence of innovations on their own past. Finally, our study also complements the work of Lee and Long (2009) who employ standardized copula distributions to model the cross-sectional dependence in asset returns in multivariate GARCH models.

Chapter 2: *Detecting Asset Price Bubbles in Real-Time through Indicator Combinations*

This chapter assesses whether emerging asset price bubbles can be detected in real-time. For this, I begin by evaluating the accuracy of existing early warning indicators for stock price bubbles. I apply these indicators to U.S. stock market data and highlight the considerable signal heterogeneity across all indicators, with frequent false positive signals during normal times and instable signals during the 1990's dot-com bubble. To improve signal accuracy, I then propose two strategies to combine signals from all individual indicators in real-time. First, I put forward a simple counting approach that requires the number of simultaneous bubble signals from all indicators to exceed a specified threshold. Second, I develop a combination indicator based on the multiple testing procedure of Romano and Wolf (2005) that controls the overall size of such a joint test. Through simulations, I show that both combination strategies provide more accurate real-time signals for the emergence and collapse of asset price bubbles than the individual indicators.

Thereby, this study contributes to a literature that aims at monitoring asset prices for the emergence of bubbles. For long, the literature has identified bubble periods as pronounced price deviations from an (HP-)filtered trend (e.g. Assenmacher-Wesche and Gerlach, 2010). Recently, however, Phillips et al. (2011, 2015) have proposed to recursively test both asset prices and their fundamental (dividend) series for explosive roots, hence extending the early work of Diba and Grossman (1988) and Evans (1991) to the real-time dimension. In this study, I show that the exact specification of the Phillips et al. (2011, 2015) indicators is crucial for their accuracy of detecting bubbles. In particular, it matters whether individual price and dividend

series or the price-to-dividend ratio are tested for explosive roots. Furthermore, I illustrate that all available individual indicators are sensitive to the number, duration and timing of bubbles in the sample. By combining individual signals, this sensitivity is markedly reduced and real-time bubble detection is considerably enhanced.

Chapter 3: *Predicting Output with Real-Time Bubble Indicators*
This chapter, based on joint work with Dirk Ulbricht, assesses whether the real-time bubble indicators developed in Chapter 2 contain valuable information for predicting real economic activity. Applied to U.S. stock and housing market data and added to an autoregressive (AR) model for output growth, we find that several bubble indicators strongly improve the forecasts from the AR model. Moreover, these bubble-augmented AR models are also highly competitive in providing accurate forecasts against a large set of 216 models based on macroeconomic and financial predictors commonly used to forecast real economic activity. In addition, we note that the best predictive bubble indicators also provide the most plausible bubble signals, providing further evidence that these indicators are capable of detecting true bubble episodes in real-time.

Our study contributes to an extensive literature that aims at identifying leading indicators for future output growth (see Stock and Watson, 2003). Next to macroeconomic variables, asset prices, and particularly stock prices, have long been employed to predict real economic variables and inflation. In univariate settings, however, their predictive value has frequently been found to be low, especially during the Great Moderation from the mid 1980's to the GFC (Stock and Watson, 2003; Rossi and Sekhposyan, 2010). In contrast, we find that asset price booms are consistently linked to higher output growth in the short-term, with a possible reversion in the long-run. Thus, our results suggest that it may be particularly the excessive boom and bust phases in asset prices that carry pronounced real economic effects. This can be explained by the presence of a stronger financial accelerator mechanism during those times, which links rising asset prices to higher investment, consumption, and, ultimately, output as suggested by Kiyotaki and Moore (1997), Holmstrom and Tirole (1997), Bernanke et al. (1999), and Martin and Ventura (2012).

Chapter 4: *Monetary Policy and Mispricing in Stock Markets*
This chapter, based on joint work with Kerstin Bernoth, investigates the role of monetary policy in misaligning stock prices from their fundamental value. Using a structural vector autoregressive (SVAR) model, we decompose the estimated response of stock prices to a monetary policy shock into a change of expected future dividends and a change in the equity risk premium. We find that only about one third of the overall impact response of stock prices can be accounted for by these two sources, which suggests a strong and systematic overreaction of stock markets to monetary policy shocks. This result lends support to the proponents of an active LATW monetary policy: By raising interest rates, the central bank can indeed lower a mispricing component in stock prices. However, this comes at the cost of dampening real economic activity and is hence only recommendable to an inflation-targeting central bank when a perceived excessive asset price boom is accompanied by economic growth and inflation above the bank's targets.

Our results challenge the theoretical predictions and empirical findings of previous literature studying the link between monetary policy and stock mispricing. In contrast to our results, Galí (2014) argues that a monetary policy tightening should, in fact, *increase* the growth rate of a *rational* asset price bubble. Galí and Gambetti (2015) provide empirical support for this hypothesis, finding that stock prices increase in response to a contractionary monetary policy shock. We show, however, that their results only hold under the restrictive identifying assumption that the central bank does not respond to exogenous changes in stock prices on impact. Instead, the findings reverse when allowing monetary policy to respond to stock price shocks instantaneously. To achieve identification of the structural shocks under this less restrictive assumption, we combine zero and sign restrictions following a recent contribution by Arias et al. (2014). While our results are at odds with the predictions of a rational bubble framework, they can be explained by mispricing arising from false subjective expectations of irrational investors as in Brunnermeier and Julliard (2008). Finally, our study complements the existing empirical literature on the effects of monetary policy on stock prices (e.g. Rigobon and Sack, 2004; Bjørnland and Leitemo, 2009) and deepens the understanding of the underlying sources for stock price changes by extending the framework of Bernanke and Kuttner (2005) to allow for the presence of a mispricing component.

CHAPTER 1

Risk Forecasting in (T)GARCH Models with Uncorrelated Dependent Innovations[1]

1.1 Introduction

In financial practice and particularly in risk management, GARCH models have become a standard econometric tool for modeling the volatility dynamics of financial returns and for ex ante prediction of risk measures such as the conditional *Value-at-Risk* (cVaR), or the conditional *Expected Shortfall* (cES).[2] Since their introduction by Engle (1982) and Bollerslev (1986), various GARCH extensions accounting for asymmetries and/or non-linear behavior in the volatility process have been made, and generally improved the understanding of the second order dynamics of financial time series. In particular the extension proposed in Glosten et al. (1993) has been widely applied for modeling stock return volatility which is known to respond stronger to bad news (i.e. negative returns) than to good news. Henceforth, we refer to this model as the *threshold GARCH* (TGARCH) model. Apart from issues related to volatility dynamics, the assumption of conditional normality in the models' standardized residuals has been questioned. Leptokurtic, non-normal conditional

[1] This chapter is based on a paper with Helmut Herwartz and Moritz Seidel which was published in *Quantitative Finance*, Vol. 17(1), 2017, 121–137, available at `http://dx.doi.org/10.1080/14697688.2016.1184303`. The views expressed in this chapter are ours and should not be attributed to the Deutsche Bundesbank or its staff. We thank Christian Conrad, Matthias Fengler, Christian Hafner, Helmut Lütkepohl, Timo Teräsvirta, and two anonymous referees, as well as participants at the *IAAE Annual Meeting 2014*, London for helpful comments and suggestions.

[2] See Angelidis and Degiannakis (2007) for an overview of GARCH cVaR- and McNeil and Frey (2000), and Zhu and Galbraith (2011) for two examples for GARCH cES forecast models.

distributions have been shown to improve both the GARCH implied approximation of empirical returns (see e.g. Bollerslev, 1990), as well as in-sample and out-of-sample forecasts of volatility and other risk measures (see Zhu and Galbraith (2011) for an overview of this literature).

In empirical work it has become a widespread convention to assume that TGARCH model innovations are identically and independently distributed (IID) with zero mean and unit variance. Hence, applied models proceed under the framework of the so-called strong GARCH form. While theoretical treatments of semi-strong GARCH models (Lee and Hansen, 1994; Escanciano, 2009; Linton et al., 2010) have shown consistency and asymptotic normality of quasi maximum likelihood (QML) estimators under weaker (mixing) properties of the underlying innovation process, the empirical scope of semi-strong GARCH models has been rarely addressed. For instance, until now it has only been suggested to implement conditional skewness (Hansen, 1994; Harvey and Siddique, 1999) and kurtosis (Brooks et al., 2005) in GARCH model innovations by conditioning their distribution on lagged returns and the estimated variance. To the best of our knowledge, however, none of the available extensions of the basic GARCH model considers the standardized residuals to exhibit an explicit serial dependence structure, i.e. a dependence of subsequent innovations in higher moments.

In this paper we uncover if TGARCH residuals are indeed IID, or show some form of higher order dependence. For the vast majority of a cross-section of 18 stock markets we find that the IID assumption is rejected by empirical innovation estimates. Instead, successive residuals from estimated univariate TGARCH models show forms of higher order dependence beyond zero correlation, and thus may carry predictive content for common risk measures. To exploit this potential predictive content, we thus aim to quantify the serial dependence structures between model innovations. For this, we propose two approaches. On the one hand, we employ *nonparametric* estimates of the conditional distributions of consecutive innovation estimates. On the other hand, we follow a suggestion in Herwartz (2013) and consider a flexible class of *standardized copula distributions* introduced to multivariate GARCH models by Lee and Long (2009). This work also relates to Chen and Fan (2006) who suggest the use of copula distributions for modeling dependence in univariate time-series and highlight their use for risk forecasting.

For a practical application, we evaluate to what degree these dependencies can be exploited for risk forecasting. In 1996, the *Basel Committee for Banking Supervision* suggested the calculation of a 1% *Value-at-Risk* (VaR) as the standard method for measuring market risk in financial institutions (Basel Committee on Banking Supervision, 2005). Moreover, to prevent financial institutions from underestimating market risks, the Basel Committee defined rules on backtesting methods for internal risk models and implemented a system of multiplication factors for penalizing those models that fail the backtesting criteria.[3] Despite its prominent use in risk management, VaR has been criticized for several drawbacks. Owing to (i) its failure in capturing the entire tail risk beyond the nominal level, and (ii) the incentives it provides for financial institutions to take on tail risks, the Basel Committee recommends to replace the VaR by the *expected shortfall* (ES) as the new standard risk measure (Basel Committee on Banking Supervision, 2016). In particular, ES refines tail risk evaluation. Moreover, to cover a larger range of extreme events it is intended to raise the nominal 1% benchmark risk level to 2.5%. These two risk levels are in the focus of our analysis.

Therefore, we conduct a large scale *out-of-sample* forecasting analysis to uncover if serial dependence in TGARCH implied model residuals can be fruitfully exploited for ex ante prediction of the cVaR and the cES. Specifically, we compare the predictive performance of nonparametric and copula based treatments of serial dependence patterns with two avenues of risk prediction that rely on the assumption of IID TGARCH innovations: Common estimates obtained from the conditionally Gaussian TGARCH model, and rival predictors where TGARCH innovations are presumed to stem from a leptokurtic IID standardized Student-t distribution. To preview some results, we find that consecutive standardized TGARCH innovations exhibit a dependence structure that differs from both conditional normality and independence, and can generally be best described either nonparametrically, or by means of standardized Clayton or Gumbel copulae with leptokurtic marginal distributions. Exploiting serial dependence patterns, cVaR and cES predictions strongly outperform risk forecasts from TGARCH models that build upon the assumption of IID innovations. This leads to more precise cVaR predictions at all risk levels, but es-

[3] When failing to employ a satisfactory internal risk model, the bank faces higher capital requirements as enforced by the multiplication factors. This reduces the funds available for lending or investing, and thereby these factors introduce an incentive to maintain a satisfactory risk model.

pecially at the 1%-level, which is most relevant according to the current rules of the Basel framework. Furthermore, cES forecasts derived from matching dependence patterns with copula distributions are shown to be more conservative than forecasts relying on the IID assumption. In line with this finding, we demonstrate that the former improve the overall forecasting accuracy, and provide sizeable economic gains for a regulating institution compared with predictions derived from quantiles of standardized Student$-t$ distributions. Overall, the forecast combination models employing copula distributions and nonparametric density estimates perform best at the relevant 1% and 2.5% nominal risk levels. Since our data set comprises times of strong market disruptions induced by the global financial crisis, we show that our method works especially well during periods of stress which need to be considered for cES prediction according to the revised Basel market risk framework.

The arguments in the remainder of this paper are illustrated by stylized features of empirical data. The next section introduces the considered cross section of 18 stock indices, and states the TGARCH(1,1,1) model that we employ for volatility estimation, extraction of model innovations and volatility forecasting. Subsequently, section 1.3 sketches a class of standardized copula distributions that we consider suitable to describe empirical serial dependence patterns of TGARCH innovations. Section 1.4 addresses the cVaR and cES forecasting in the framework of the proposed model class, nonparametric estimation of conditional distributions, and the applied backtesting procedures. Section 1.5 discusses our results. Section 1.6 concludes.

1.2 Dependence in GARCH innovations

In this section we introduce the cross section of stock markets and state the TGARCH model. We then argue that the typical assumption of IID innovations is likely at odds with stylized facts of empirical data.

1.2.1 Data and the TGARCH model

We analyze return patterns for a cross section of 18 stock market indices. For each of these indices about 2,300 daily observations cover the time period from May 1st,

2003 until April 30th, 2012.[4] The TGARCH(1,1,1) model introduced by Glosten et al. (1993) accounts for asymmetries in the volatility response and is used as a filter to obtain the implied model innovations.[5] Conditional on the information available at time $t-1$, denoted Ω_{t-1}, first and second order return characteristics are given as

$$r_t = \mu_t + \sigma_t \xi_t, \tag{1.1}$$

$$\sigma_t^2 = \gamma_0 + \beta_1 \sigma_{t-1}^2 + \gamma_1 (r_{t-1} - \mu_{t-1})^2 + \gamma_1^- (r_{t-1} - \mu_{t-1})^2 I(r_{t-1} - \mu_{t-1} < 0),\ t = 1, \ldots, T, \tag{1.2}$$

where $I(.)$ is an indicator function, and T indicates the available sample size conditional on presample values. Moreover, σ_t^2 is the conditional variance of the daily log-returns r_t, μ_t is the conditional mean, and γ_0, γ_1, γ_1^- and β_1 denote the variance response parameters. Focusing on the analysis of daily stock returns, we follow common practice and assume $\mu_t = 0$ in (1.1) and, thus, $r_t = \sigma_t \xi_t$.

The TGARCH-model innovations ξ_t in (1.1) are typically assumed to be IID. For purposes of (Q)ML estimation the Gaussian assumption is common, i.e.

$$\xi_t = r_t / \sigma_t \overset{iid}{\sim} N(0,1). \tag{1.3}$$

The IID assumption in (1.3) embeds r_t within the class of strong GARCH processes. In the class of semi-strong GARCH models that we will employ, ξ_t is required to be strictly stationary and ergodic such that $E[\xi_t|\Omega_{t-1}] = 0$ and $\text{Var}[\xi_t|\Omega_{t-1}] = 1$ (*scaled martingale difference sequence, sMDS*).[6] Hamadeh and Zakoïan (2011) justify QML estimation of semi-strong TGARCH models presuming that model innovations take the form of an sMDS process.

[4] Stock price data have been obtained from Yahoo!Finance (finance.yahoo.com) except for the DJIA 30 which has been drawn from the Federal Reserve Bank of St. Louis, Economic Research (http://research.stlouisfed.org/fred2/).

[5] Based on the conditional standard deviation σ_t, Zakoian (1994) proposed a similar threshold GARCH specification.

[6] The distinction of strong, semi-strong and weak forms of the GARCH model has become prominent in the framework of temporal aggregation of GARCH processes. Drost and Nijman (1993) have shown that only the weak form of GARCH models is closed under temporal aggregation.

In this work we argue that consecutive empirical (T)GARCH innovations are at odds with the typically imposed IID assumption.[7] Combined with the assumption $\mu_t = 0$, the TGARCH model can be regarded as a filter approach to isolate model implied innovations. In consequence, filtered innovations may show remaining dependence from neglecting non-zero conditional expectations ($\mu_t \neq 0$) or applying a false/inappropriate volatility model. Against this background, we provide statistical evidence on the full sample (Table 1.B.1) and on subsamples (Tables 1.B.2 and 1.B.3) showing that the suggested filter ($\mu_t = 0$, TGARCH(1,1,1)) is sufficient to obtain innovations which are in line with the typical first and second order assumptions.

The set of analyzed stock market indices and full sample parameter estimates for the TGARCH(1,1,1) models are shown in Table 1.B.1. Significant *leverage effects* ($\gamma_1^- > 0$) can be diagnosed for all markets. In consequence, it appears justified to concentrate on this particular variant of conditionally heteroskedastic processes. Further, the parsimonious specification of just one GARCH and ARCH term is supported by the BIC criterion for all markets (not shown). Table 1.B.1 also documents diagnostics testing the TGARCH(1,1,1) residuals against *serial correlation* by means of the LB statistic of Ljung and Box (1978) and against *remaining heteroskedasticity* by means of the Lagrange Multiplier (LM) statistic of Lundbergh and Teräsvirta (2002).[8] The potential of serial correlation can be neglected, since p-values of the LB statistics with 20 lags (LB(20)) are well above five percent for almost all markets. While for some markets the LM statistics appear to hint at some degree of remaining heteroskedasticity with five percent significance, we notice that extending the TGARCH model orders we do not obtain more favorable diagnostic outcomes. These results also hold for a rolling window subsample analysis of 1,000 days repeated at every 100 days of the sample (see Tables 1.B.2 and 1.B.3).[9] We conclude that the TGARCH(1,1,1) is largely sufficient to remove dependence in the residuals' first and second order moments, but acknowledge that in individual cases higher lag orders or other models might provide an even better filter.

[7] A natural extension of our work is to evaluate higher order dependence of innovations at even longer lags. Addressing such dependence structures through our proposed modeling approach, however, increases the computational demands exponentially. We hence restrict the analysis to the bivariate case of *consecutive* innovations.

[8] While the applied LM statistic is robust against non-Gaussian innovations ξ_t, potential outliers of standardized returns might give rise to biased LM statistics.

[9] We further note that the TGARCH model is selected over the GARCH model by the BIC at all iterations except for the SSEC index.

1.2.2 Higher-order dependence of uncorrelated innovations

As a first indication of higher-order dependence of consecutive innovations, Figure 1.2.1 displays scatter diagrams of the $T > 2{,}000$ estimated model residuals $\hat{\xi}_t$ and $\hat{\xi}_{t+1}$ for four stock market indices. From the clustering of observations in the bottom left corners (upper right for NIKKEI), it appears that the assumptions of both a joint normal distribution and independence are questionable. In contrast to an IID setting, there seems to prevail a remaining dependence structure characterized by a clustering of subsequently negative (positive for NIKKEI) innovations. Hence, empirical TGARCH residuals appear to be characterized by joint lower or upper *tail dependence*.

Following the graphical insights from Figure 1.2.1, we continue with explicitly testing for the higher moment dependence. For this purpose we concentrate on means of third and fourth order co-moments determined for the vectors of cross-products $\upsilon_t = \left((\hat{\xi}_{t+1}^2\hat{\xi}_t), (\hat{\xi}_{t+1}\hat{\xi}_t^2)\right)'$ and $\psi_t = \left((\hat{\xi}_{t+1}^2\hat{\xi}_t^2 - 1), (\hat{\xi}_{t+1}^3\hat{\xi}_t), (\hat{\xi}_{t+1}\hat{\xi}_t^3)\right)'$, $t = 1, 2, \ldots, T-1$.[10] Note that in case of IID innovations one has $E[\upsilon_t] = 0_2$ and $E[\psi_t] = 0_3$, where 0_J is a J-dimensional zero vector. Patterns of dependence of consecutive innovations may show up in third or fourth order co-moments such that $E[\upsilon_t] \neq 0_2$ or $E[\psi_t] \neq 0_3$. We test the null hypotheses $E[\upsilon_t] = 0_2$ and $E[\psi_t] = 0_3$ by means of the following *moment based test statistics*, respectively,

$$\lambda_{iid}^{(3)} = T\bar{\upsilon}' \left(\mathrm{Cov}[\upsilon_t]\right)^{-1} \bar{\upsilon} \xrightarrow{d} \chi^2(2) \text{ and } \lambda_{iid}^{(4)} = T\bar{\psi}' \left(\mathrm{Cov}[\psi_t]\right)^{-1} \bar{\psi} \xrightarrow{d} \chi^2(3), \qquad (1.4)$$

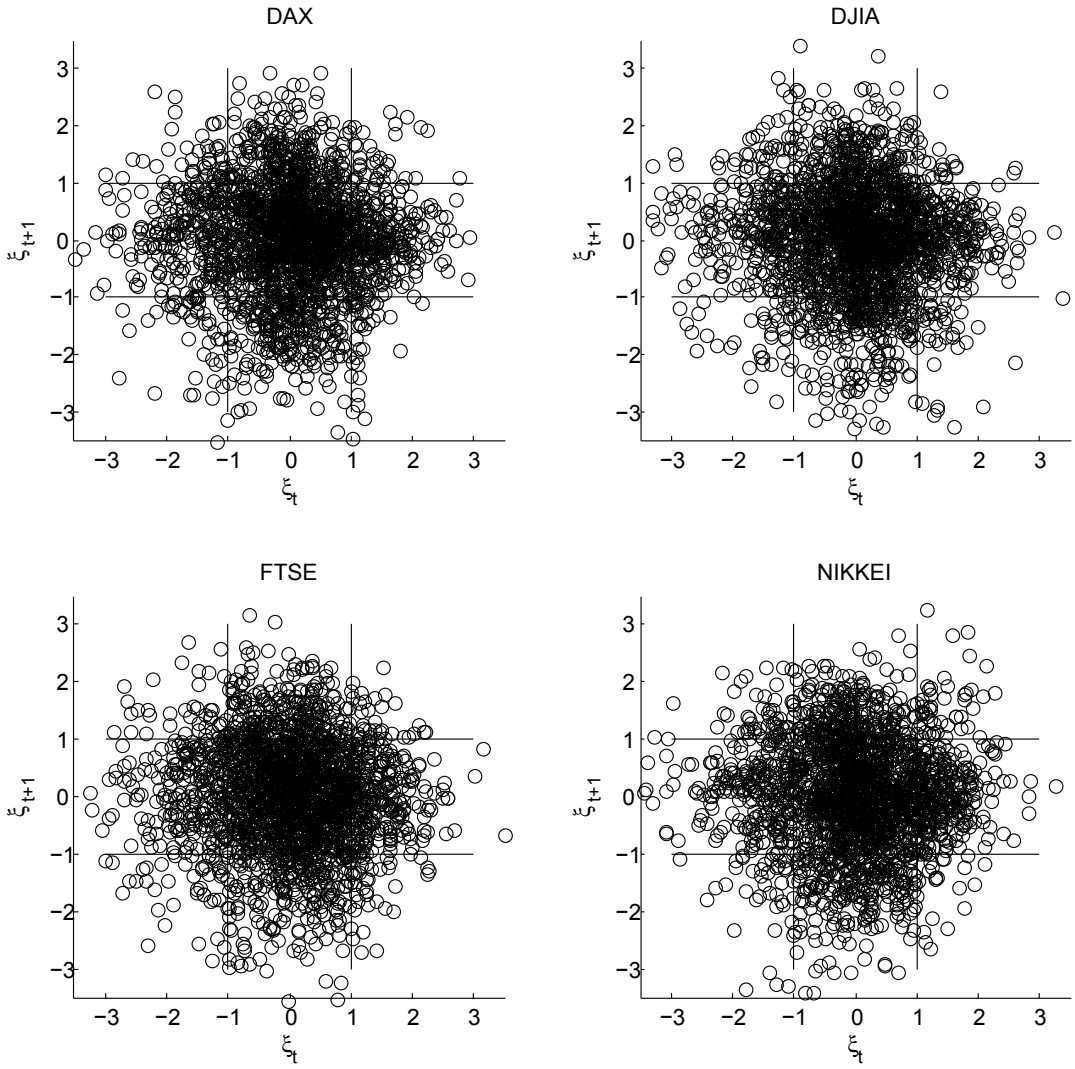
where $\bar{\upsilon} = (1/T)\sum_{t=1}^{T} \upsilon_t$ and $\bar{\psi} = (1/T)\sum_{t=1}^{T} \psi_t$. Covariances $\mathrm{Cov}[\upsilon_t]$ and $\mathrm{Cov}[\psi_t]$ are determined by means of Monte Carlo techniques using simulated sequences of Gaussian innovations.[11]

Table 1.B.4 documents test statistics $\lambda_{iid}^{(3)}$ and $\lambda_{iid}^{(4)}$ along with respective p-values. Overall, empirical estimates of TGARCH innovations hardly accord with the commonly held IID assumption. At the five percent significance level, we find that nine

[10] Being aware of more general tests of the null hypothesis of independence, as for instance, the test by Broock et al. (1996))(BDS test), we concentrate on moment based statistics that directly target dependence among consecutive innovations.

[11] To determine the covariance of υ_t and ψ_t we draw 5,000 replications of Gaussian processes of length T. For the simulation of leptokurtic innovation distributions we draw IID processes of Student$-t$ distributed random variables with ten degrees of freedom which are standardized by their empirical standard deviations.

Figure 1.2.1: Standardized TGARCH(1,1,1) residuals (ξ_t vs. ξ_{t+1}) for four indices



Notes: The figure shows the standardized TGARCH(1,1,1) residuals (ξ_t vs. ξ_{t+1}) for four indices from QML estimation. Solid lines indicate the absolute unit levels.

and 14 stock market indices (out of 18) are characterized by third and fourth order dependence, respectively.

In empirical work, TGARCH innovations are often found to be characterized by excess kurtosis. Thus, we also estimate $\mathrm{Cov}[\upsilon_t]$ and $\mathrm{Cov}[\psi_t]$ by means of simulated standardized Student $t-$distributed random variables with ten degrees of freedom. Moreover, the determination of higher order co-moments by means of empirical TGARCH innovations may be flawed by potential outliers. To account for such

effects we also determine the co-moments entering the dependence statistics in (1.4) from trimmed innovations. For the trimming we replace estimates $\hat{\xi}_t$ which are above (below) the 99% (1%) quantile of the unconditional distribution by these respective quantiles. Along these lines and using standardized $t(10)$ innovations for Monte Carlo (MC) diagnostics, third and fourth order dependence remains as a significant data characteristic in five and 17 markets, respectively. In comparison with patterns of coskewness, fourth order codependence appears to be more evident for empirical TGARCH innovations. Hence, we focus on symmetric marginal distributions in the following analysis.[12]

In summary, graphical displays and diagnostic tests highlight that empirical TGARCH innovations are likely at odds with the commonly held IID assumption. Thus, in terms of the involved conditional and unconditional distribution functions (CDFs) we likely have $F(\xi_{t+1}|\xi_t) \neq F(\xi_{t+1})$. With this empirical diagnosis at hand, it is of immediate interest if such dependence patterns can be exploited to improve common TGARCH based predictions of conditional risk. To model serial dependence of TGARCH innovations we pursue two strategies. On the one hand we design a huge class of structured parametric models, and on the other hand we follow an unstructured nonparametric approach.

1.3 The copula-TGARCH model

While the dependence in higher order co-moments might be easily diagnosed by means of dependence diagnostics $\lambda^{(3)}_{iid}$ and $\lambda^{(4)}_{iid}$ in (1.4), these statistics are not informative about how to arrive at (parametric) conditional ex ante risk predictors. Furthermore, the heterogeneity of empirical co-moments displayed in Table 1.B.4 (third order co-moments of either sign, distinct significance patterns for third and fourth order co-moments) raises doubts that a full parametric framework can indeed accurately describe all forms of higher order dependence. Therefore, we opt for a flexible semi-parametric modeling framework that allows to capture varying forms of higher order dependence and to predict the conditional risk in asset returns.

[12] This result also holds in the sub-sample analysis. Evidence for coskewness in sub-samples is indicated for seven markets, while strong codependence in fourth order moments is present in twelve markets. Extending lag orders of the TGARCH specification does not remove these dependence patterns.

In this section we outline a model class that has been advocated in Herwartz (2013) to improve the performance of linear autoregressions in macroeconomic forecasting. In particular, we employ the model class of standardized copula distributions which allow a flexible and data driven evaluation of nonlinear relationships and, thus, the embedding of general dependence patterns $F(\xi_{t+1}|\xi_t) \neq F(\xi_{t+1})$. For this, we proceed in two steps. First, we outline the candidate standardized copula distributions used to describe empirical dependence patterns in the data and to improve risk forecasts. We also describe how to draw samples from these distributions that fulfill the sMDS property as required for TGARCH innovations. Second, we then sketch a data driven adaptive approach to select a particular distribution from this large model class that describes the empirical dependencies best. Monte Carlo techniques prove useful to quantify model implied serial dependence patterns and to match particular model specifications to empirical data.

1.3.1 Standardized copula implied dependence structures

1.3.1.1 Standardized copula distributions

A priori the space of potential serial dependence patterns appears overly large to define a particular parametric model nesting the most promising nonlinear model approaches. Instead, we opt for a modeling approach that allows to flexibly translate heterogeneous empirical dependence structures into the two dimensional space of (ξ_{t+1}, ξ_t). The standardized copula distributions introduced by Lee and Long (2009) embed a such rich dependence structure. Before we specify the copula distributions considered in the empirical analysis, we describe their general form and how we can obtain random samples that meet the sMDS requirements. Henceforth $\mathbf{v}_t = (\xi_{t+1}, \xi_t)'$ denotes a vector of consecutive TGARCH innovations which fulfill the sMDS property with unit variance by construction, i.e. $E[\xi_{t+1}|\Omega_t] = 0$ and $\mathrm{Var}[\xi_{t+1}|\Omega_t] = 1$. To draw samples of $\mathbf{v}_t$, we proceed along the following three steps:

1. Generate serially correlated processes $\mathbf{w}$ from parametric copulae:
   Let

$$\mathbf{w} \sim \mathcal{C}_\delta(\mathcal{G}_1(w_1), \mathcal{G}_2(w_2), \ldots, \mathcal{G}_{T+1}(w_{T+1})),\ \mathrm{E}[\mathbf{w}] = 0,\ \mathrm{Cov}[\mathbf{w}] = \Sigma, \qquad (1.5)$$

denote a $(T+1)$-dimensional vector of copula distributed random variables where $\mathcal{G}_t(.)$, and $\mathcal{C}_\delta(.)$ denote marginal distribution functions and a copula function with parameter δ that controls the strength of dependence, respectively.[13] The covariance matrix Σ in (1.5) is positive definite and its diagonal elements (marginal variances) are unity.

For large T, random vectors $\mathbf{w}$ can be composed by drawing from conditional copula distributions. For this purpose we sequentially draw serially dependent realizations w_{t+1} of the random innovations from the conditional copula distributions implied by (1.5) conditional on w_t that is available from the t-th draw. We implement this by transformations of the p-values u_t and u_{t+1} to w_t and w_{t+1} respectively, according to the marginal distribution, e.g. $w_t = \mathcal{G}_t^{-1}(u_t)$. The procedure is initiated by drawing u_{-1000} from the uniform distribution. To immunize the simulation chain from initial conditions, we discard the first 1,000 draws.[14] Removing serial correlation from $\{w_t\}_{t=1}^T$ requires a consistent estimate for the covariance matrix Σ. Therefore, we draw S = 1,000 samples of $\mathbf{w}$, each of size T = 2,000 and use the average empirical covariance matrix estimate as an approximation of Σ. With sufficiently large S the covariance structure implied by a particular distributional model is estimated consistently.[15]

2. Extract serially uncorrelated innovations:
   Let C denote a lower triangular *Cholesky factor* such that $\Sigma = CC'$. Then, the coefficients on and left from the diagonal of C provide a weighting scheme to transform first order correlated data w_t in $\mathbf{w}$ to serially uncorrelated variables

[13] Copulae have become an established framework to model dependence structures of vector valued random variables (see Cherubini et al. (2004) or Nelsen (2006) for textbook treatments). Recently, Chen and Fan (2006) have proposed a copula approach to formalize general dependence patterns of successive univariate time series variables.

[14] Sampling from conditional copulae follows Cherubini et al. (2004), and is accomplished by means of the Matlab function *'copularnd'* adapted to draw only u_{t+1}, conditional on u_t, $t = -1000, \ldots, T$.

[15] Increasing the number of draws S beyond 1,000 does not change the approximation of bivariate distributions as described below in section 1.3.2.

$\widetilde{w}_t$. For instance, from the last row of C we get the following weighting scheme

$$\widetilde{w}_t = \sum_{i=0}^{T_0} c_{T,T-i} w_{t-i}, \tag{1.6}$$

where $T - T_0$ is some cut-off threshold, and $c_{T,T-i}$ indicate typical elements in the last row of C. By construction, the process $\widetilde{w}_t$ is serially uncorrelated but is likely at odds with the sMDS assumption made for innovations in GARCH specifications of the semi-strong form.

3. Convert to a sMDS sequence:
   To convert the series $\widetilde{w}_t$ to a sMDS sequence, we employ two regression designs to determine the expectation of either $\widetilde{w}_t$ or $\widetilde{w}_t^2$ conditional on most recent time series information. Specifically, the following regressions are estimated for each of the $S = 1{,}000$ samples with $T = 2{,}000$ observations:

$$\widetilde{w}_t = \beta_{1,0} + \sum_{i=1}^{5}\sum_{j=1}^{4} \beta_{1,i,j} \widetilde{w}_{t-i}^j + \omega_{1,t}, \tag{1.7}$$

$$\widetilde{w}_t^2 = \beta_{2,0} + \sum_{i=1}^{5}\sum_{j=1}^{4} \beta_{2,i,j} \widetilde{w}_{t-i}^j + \omega_{2,t}. \tag{1.8}$$

   Using the estimated parameters $\hat{\beta}_{k,i,j}$ averaged across all samples S, we obtain the conditional expectations $\widehat{m}_t = E[\widetilde{w}_t|\Omega_{t-1}]$ and $\hat{s}_t = E[\widetilde{w}_t^2|\Omega_{t-1}]$ from the auxiliary predictive regressions (1.7) and (1.8). Then, innovation sequences that accord with the martingale assumption read as

$$\xi_t = \frac{\widetilde{w}_t - \widehat{m}_t}{\sqrt{\hat{s}_t - \widehat{m}_t^2}}, \; t = 1, 2, \ldots, T+1. \tag{1.9}$$

   From the random variables defined in (1.9) we determine bivariate tuples that are considered to comprise consecutive GARCH innovations, i.e. $\mathbf{v}_t = (\xi_{t+1}, \xi_t)'$.[16]

[16] We find that the obtained ξ_t fulfill the requirements $E[\xi_{t+1}|\Omega_t] = 0$ and $\text{Var}[\xi_{t+1}|\Omega_t] = 1$. In addition, we find correlation close to one between ξ_t and $\widetilde{w}_t$. Hence, we conclude that the $\widetilde{w}_t$ show only minor deviations from sMDS patterns already.

While the conditional first and second order moments of elements in $\mathbf{v}_t$ are identical over alternative copulae, their conditional quantiles are model-specific, since each copula is characterized by a distinct conditional CDF. Depending on the actual choice of $\mathcal{G}$, $\mathcal{C}_\delta$, δ and rotations of $\mathcal{C}_\delta$, the above procedure allows to obtain a large set of standardized copula distributions describing a flexible class of conditional distributions $F_j(\xi_{t+1}|\xi_t)$, $j = 1, 2, \ldots, J$.

1.3.1.2 A model class based on standardized copula distributions

The set of copulae employed in this study consists of three *Archimedean* families, namely the *Clayton* ($\mathcal{C}_\delta^C$), *Gumbel* ($\mathcal{C}_\delta^G$) and *Frank* ($\mathcal{C}_\delta^F$) copula. With arguments $u_t = \mathcal{G}_t(w_t)$, these read, respectively, as

$$\mathcal{C}_\delta^C(u_t, u_{t+1}) = \max\left[(u_t^{-\delta} + u_{t+1}^{-\delta} - 1)^{-\frac{1}{\delta}}, 0\right], \ \delta > 0, \tag{1.10}$$

$$\mathcal{C}_\delta^G(u_t, u_{t+1}) = \exp\left(-\left[(-\ln u_t)^\delta + (-\ln u_{t+1})^\delta\right]^{\frac{1}{\delta}}\right), \ \delta > 1, \text{ and} \tag{1.11}$$

$$\mathcal{C}_\delta^F(u_t, u_{t+1}) = -\frac{1}{\delta}\ln\left(1 + \frac{(\exp(-\delta u_t) - 1)(\exp(-\delta u_{t+1}) - 1)}{\exp(-\delta) - 1}\right), \ \delta > 0. \tag{1.12}$$

In addition to the specifications in (1.10) to (1.12), we consider one rotation of each copula to increase the space of possible tail dependencies. Random variables from rotated copulae are obtained by replacing u_{t+1} by $1 - u_{t+1}$ in (1.10)–(1.12). The rotations in part revert the dependence patterns of their unrotated counterparts.[17]

To further improve the flexibility in matching residual sequences with copula innovations, we enlarge the model space along two additional dimensions. First, we consider a set of parameters δ for each (rotated) copula such that, prior to standardization, the strength of dependence (measured by means of *Kendall's* τ) is controlled. The target values for τ are $\tau^\bullet = \{0.010, 0.020, ..., 0.060, 0.075, 0.100, 0.150, ..., 0.750\}$.[18] Second, we allow for five alternative marginal distributions. Since the marginals in (1.10)–(1.12) refer to sequential draws from a univariate

[17] Figures 1.A.3 and 1.A.4 in the appendix show exemplary scatter diagrams for random draws of the bivariate Gaussian and Archimedean copulae (Figure 1.A.3: unrotated, Figure 1.A.4: rotated). The rotated Clayton copula and the unrotated Frank and Gumbel copulae are characterized by patterns of lower tail dependence. It can consequently be argued that these specifications likely provide the best distributional description of TGARCH residuals for the stock indices displayed in Figure 1.2.1.

[18] For the relation between δ and τ of the three copula families, see Cherubini et al. (2004).

uncorrelated innovation process, we presume $\mathcal{G}_t(.) = \mathcal{G}_{t+1}(.)$. Allowing for distinguished degrees of conditional leptokurtosis, we consider marginals stemming from the Gaussian or standardized Student t-distributions with $\nu = 5, 10, 15, 20$ degrees of freedom. Note that for the empirical analysis of daily stock returns estimated innovation processes are often characterized by degrees of freedom between five and twelve (see, e.g., Table 1.B.1). All in all, distinguishing between alternative degrees of dependence, rotations and marginal distributions for the Clayton, Frank and Gumbel family, we obtain $J = 635$ distinct distributional models, including the IID Gaussian and standardized Student-t distributions of IID TGARCH innovations.

1.3.2 Copula matching

In light of a large class of candidate models $F_j(\xi_{t+1}|\xi_t)$, $j = 1, 2, \ldots, J$, model choice is an essential step to exploit implied higher order dependence patterns of ξ_t for risk prediction. We refrain from direct estimation of the model parameters of the TGARCH model in (1.1) and (1.2) jointly with the dependence parameter of the Copula distribution in (1.5) for the following reasons: First, to our knowledge (non-linear) maximum likelihood as proposed by Lee and Long (2009) or a multi-step estimation procedure has not been established theoretically in the present framework of serially dependent, sMDS model innovations. Second, in our iterative out-of-sample forecasting applications, a repeated assessment of TGARCH likelihoods at each point in the sample would imply an overly large computational burden. Third, the model space outlined above is very large such that each model may suffer from misspecification in local time windows. Finally, to allow for structural variations one may a priori opt for an estimation/calibration approach that can be most easily implemented adaptively.

To determine the best distributional description of the estimated model residuals by a particular standardized copula distribution, denoted $F_{j^*}(\xi_{t+1}|\xi_t)$, we contrast their discretized empirical distribution against the counterparts of simulated copulae innovations by means of a *distance measure*. This reads as

$$D^{(j)} = \sum_m \sum_n \frac{\left(H^{(j)}_{(m,n)} - \hat{h}_{(m,n)}\right)^2}{H^{(j)}_{(m,n)}}, \tag{1.13}$$

where $\hat{h}_{(m,n)}$ and $H^{(j)}_{(m,n)}$ are two-dimensional frequency estimates characterizing the empirical distribution of consecutive empirical TGARCH innovations and a particular simulated standardized copula distribution, respectively. To determine $\hat{h}_{(m,n)}$ and $H^{(j)}_{(m,n)}$ we employ a bivariate grid consisting of $12 \times 12 = 144$ cells. Specifically, $\hat{h}_{(m,n)}$ is estimated as

$$\hat{h}_{(m,n)} = \frac{1}{T-1} \sum_{t=1}^{T-1} I((a_m \leq \hat{\xi}_{t+1} < a_{m+1})(a_n \leq \hat{\xi}_t < a_{n+1})), \tag{1.14}$$

where $\hat{\xi}_t$ are the standardized TGARCH innovations, $\hat{\xi}_t = r_t/\hat{\sigma}_t$, and a_m, a_n, are from the grid vector $a = (-1000, -2.5, -2.0, \ldots, 2.5, 1000)'$ and $m, n = 1, \ldots, 12$.

To obtain a reliable approximation of the distribution of selected copulae, $S =$ 1,000 samples of length $T =$ 2,000 are drawn from each parameterized (rotated) copula, and $H^{(j)}_{(m,n)}$ is determined as an average estimate

$$H^{(j)}_{(m,n)} = \frac{1}{S} \sum_{s=1}^{S} h^{(j,s)}_{(m,n)}, \tag{1.15}$$

$$\text{with } h^{(j,s)}_{(m,n)} = \frac{1}{T-1} \sum_{t=1}^{T-1} I((a_m \leq \xi^{(j,s)}_t < a_{m+1})(a_n \leq \xi^{(j,s)}_{t+1} < a_{n+1})).$$

The minimization of $D^{(j)}$ in (1.13) obtains a particular member of the class of standardized and discretized copula distributions $H^{(j^*)}$, $j^* = \arg\min_j D^{(j)}$, which corresponds to a respective distribution $F_{j^*}(\xi_{t+1}|\xi_t)$. This distribution is employed in the out-of-sample forecasting exercises to determine risk predictors. Illustrating the potential of the employed model class in approximating the distributional patterns of empirical stock return indices, Table 1.B.5 documents in-sample distance statistics for the cross section of estimated TGARCH innovation processes.

From Table 1.B.5 we observe the common result that under an IID setting leptokurtic distributions (standardized Student-t) offer a closer fit to empirical data in comparison with the Gaussian distribution. In addition, for each market there is at least one and generally several dependent distributions that outperform the IID patterns markedly. In particular, these are the Clayton copula with $t(10)$ marginal distribution and the Gumbel copula with $t(5)$ and $t(10)$ marginals which fit the data best in eight, two, and three cases, respectively. Overall, a member of the

Clayton family provides the best fit in ten cases, while the Gumbel fits best for seven markets.[19] Supporting the selection of the marginal distributions, the results in Table 1.B.5 closely mirror the estimated degrees of freedom as documented in Table 1.B.1. Moreover, Figure 1.3.2 shows distance measures between the best fitting rotated Clayton copula (with $t(10)$ marginals) and TGARCH(1,1,1) innovations for four stock market indices.[20] For each market this measure allows for a unique identification of the distance-minimizing distribution. In summary, we conclude that the (rotated) Clayton copula obtains most accurate approximations of the serial dependence structure that underlies consecutive TGARCH innovations.

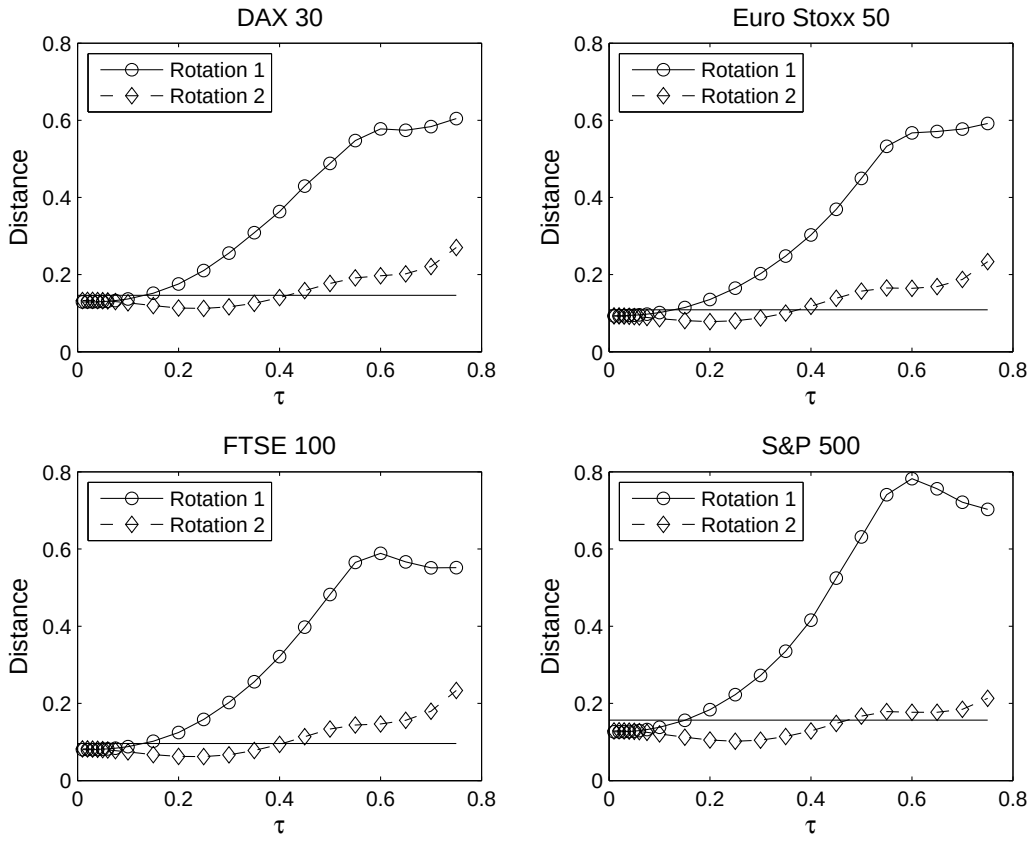
1.4 A large scale comparison of risk predictors

The previous section has shown that standardized copula distributions offer a promising framework to flexibly capture the higher order dependence structure between consecutive TGARCH model innovations. Thus, we now assess whether these dependence structures can be fruitfully exploited for one-step pseudo-out-of-sample forecasts of popular risk measures for asset returns. Therefore, this section formally describes the conditional one-step ahead ex ante risk measures (cVaR and cES) and the evaluation criteria that will be of core interest for the empirical analysis. To disentangle in-sample and (pseudo) out-of-sample time series observations, time indices t and forecast origins τ are distinguished henceforth. Moreover, Ω_τ is shorthand to summarize the time series information available in time τ, $\tau = W, W+1, \ldots, T-1$, and T is the total number of return quotes per market (see Table 1.B.1). With a slight abuse of the (in-sample) time indexation in (1.2), the sample information used to determine risk forecasts consists of stock returns $\{r_t\}_{t=\tau-W+1}^{\tau}$, where W is the length of (rolling) estimation windows. At each forecast origin τ, the TGARCH(1,1,1) model is re-estimated, and one-step ahead risk forecasts for the cVaR and the cES are made.

[19] As expected from Figures 1.A.3 and 1.A.4, we find that the best fits are generally provided by the rotated Clayton copula with a medium dependence parameter τ of about 0.20 to 0.25. For Gumbel, the unrotated copula is preferred with stronger dependence τ of 0.55 to 0.60. These results are stable across markets.

[20] The empirical findings are robust for all stock markets and the four indices have been selected for illustrative purposes only.

Figure 1.3.2: Distance between discretized distributions from the rotated Clayton copula (with $t(10)$ marginals) and of TGARCH(1,1,1) residuals for four stock indices



Notes: The figure shows the distance between discretized distributions from the Clayton copula (with $t(10)$ marginals) across different dependence parameters τ and of TGARCH(1,1,1) residuals for four stock indices. 'Rotation 1' refers to the 'unrotated' copula innovations, whereas 'Rotation 2' refers to the copula where u_{t+1} has been replaced by $1-u_{t+1}$ (see equations (1.10)-(1.12)). The solid horizontal line indicates the distance between the Gaussian and residuals' distributions.

For this, we employ seven rival models: First, we provide benchmark forecasts from two models that rely on the assumption of IID model innovations following a Gaussian or leptokurtic standardized Student$-t$ distribution. Second, we employ two predictive models that aim to exploit the dependence between consecutive model residuals either by means of a nonparametric estimation of their distribution or by fitting their distribution to one of the copula counterparts from the large set of candidate distributions $F_j(\xi_{t+1}|\xi_t)$, $j = 1, 2, \ldots, J$. Finally, we also assess the performance of three forecast combinations of the above described four individual

forecasts. The nonparametric estimation and the copula matching are repeated at each forecast origin after re-estimating the TGARCH. By this, we also implicitly allow for time-varying distributions and dependence patterns.

1.4.1 Ex ante risk measures

Conditional on Ω_τ, the cVaR at level α is defined as the (negative) quantile q_α for which the probability of the one period-ahead return being less than q_α is at most α. We distinguish nominal *coverage levels* $\alpha = 0.01, 0.025$ and $\alpha = 0.05$. Formally, a first order approximation of the cVaR reads as

$$\begin{aligned} \text{cVaR}^\bullet_\alpha(r_{\tau+1}|\Omega_\tau) &= -F^{-1}_{\bullet,\alpha}(r_{\tau+1}|\Omega_\tau) = -q^\bullet_\alpha(r_{\tau+1}|\Omega_\tau) \\ &= -\hat{\sigma}_{\tau+1} q^\bullet_\alpha(\xi_{\tau+1}|\Omega_\tau) \approx -\hat{\sigma}_{\tau+1} q^\bullet_\alpha(\xi_{\tau+1}|\xi_\tau), \end{aligned} \tag{1.16}$$

where $F^{-1}_{\bullet,\alpha}(r_{\tau+1}|\Omega_\tau)$ denotes the inverse of the conditional CDF (cCDF) of returns $r_{\tau+1}$. Henceforth, '$\bullet$' is used to indicate a particular model approach to select or determine the cCDF $F_\bullet(r_{\tau+1}|\Omega_\tau)$ for the nonparametric model or the best-fitting copula. Related to the cVaR, an approximation of the cES of $r_{\tau+1}$ is

$$\begin{aligned} \text{cES}^\bullet_\alpha(r_{\tau+1}|\Omega_\tau) &= -\text{E}_\bullet\left[r_{\tau+1}|\left(r_{\tau+1} < -\text{cVaR}^\bullet_\alpha(r_{\tau+1}|\Omega_\tau)\right), \Omega_\tau\right] \\ &\approx -\hat{\sigma}_{\tau+1}\text{E}_\bullet[\xi_{\tau+1}|(\xi_{\tau+1} < q^\bullet_\alpha(\xi_{\tau+1}|\xi_\tau))]. \end{aligned} \tag{1.17}$$

By construction, the cES exceeds the cVaR. Determining either the cVaR or the cES requires an estimate of the cCDF of returns $F_\bullet(r_{\tau+1}|\Omega_\tau)$. Since $r_{\tau+1} = \sigma_{\tau+1}\xi_{\tau+1}$, this estimate depends on both the (one-period ahead) prediction of the conditional variance $\sigma_{\tau+1}$ and on the cCDF $F_\bullet(\xi_{\tau+1}|\Omega_\tau) \approx F_\bullet(\xi_{\tau+1}|\xi_\tau)$.[21] A one-period ahead volatility forecast for the TGARCH(1,1,1) can be obtained from (1.2) by means of QML parameter estimates,

$$\hat{\sigma}^2_{\tau+1} = \hat{\gamma}_0 + \hat{\gamma}_1 r^2_\tau + \hat{\gamma}^-_1 r^2_\tau I(r_\tau < 0) + \hat{\beta}_1 \hat{\sigma}^2_\tau. \tag{1.18}$$

[21] Graphical displays of empirical innovations ξ_{t+1} vs. ξ_{t-1} (or respective higher order co-moments) illustrate very minor (if any) higher order dependence among filtered TGARCH innovations. While the direct dependence is already weak, the approximation $q^\bullet_\alpha(\xi_{\tau+1}|\Omega_\tau) \approx q^\bullet_\alpha(\xi_{\tau+1}|\xi_\tau)$ builds upon negligibility of historic information $\xi_{\tau-1}, \xi_{\tau-2}, \ldots$ given that ξ_τ enters the formation of the conditional quantile.

Conditional on a unique volatility predictor $\hat{\sigma}^2_{\tau+1}$, we then use four competing modeling choices of $F_\bullet(\xi_{\tau+1}|\xi_\tau)$. The first two of these predictors rely on the IID assumption of Gaussian or standardized Student$-t$ distributions that is commonly made for the TGARCH innovations.[22] These two forecast models provide the IID benchmarks against which we evaluate our proposed predictors that exploit potential serial dependence of TGARCH innovations. Here, we first estimate the conditional distribution without imposing any parametric structure ($F_{np}(\xi_{\tau+1}|\xi_\tau)$). Second, we select a particular candidate from the class of standardized copula distributions as described in section 1.3.2, and exploit the respective conditional distribution $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ for risk prediction.

1.4.2 Estimation of conditional CDFs

Since the nonparametric and copula-based predictors for (1.16) and (1.17) require the inversion of the model innovations' cCDF, we employ a Nadaraya-Watson (NW) estimator to approximate the distribution of $\xi_{\tau+1}$ given ξ_τ from the estimated TGARCH innovations. To outline the NW cCDF estimate we adopt the common convention to indicate with y and x the dependent $(\xi_{\tau+1}, \xi_{t+1})$ and conditioning variable (ξ_τ, ξ_t), respectively. The nonparametric cCDF $F_{np}(\xi_{\tau+1}|\xi_\tau)$ is evaluated along equidistant grid points $y \in \boldsymbol{y}$ and $x \in \boldsymbol{x}$ with $\boldsymbol{y} = (-4.0, -3.9, \ldots, 4.0)'$ and $\boldsymbol{x} = (-4.0, -3.9, \ldots, 4.0)'$. Using the sample information from rolling windows, $\{\xi_t\}_{t=\tau-W}^{\tau}$, the cCDF estimate reads as

$$\hat{F}_{np}(y|x) = \frac{\sum_{t=\tau-W}^{\tau-1} K_h(x - \xi_t) I(\xi_{t+1} \leq y)}{\sum_{t=\tau-W}^{\tau-1} K_h(x - \xi_t)}, \tag{1.19}$$

where $K_h(u) = \frac{1}{h} K(u/h)$ is a kernel function and $h > 0$ is the bandwidth parameter.[23] From the discretized estimate $\hat{F}_{np}(y|x)$, nonparametric cVaR and cES estimates can be obtained. The conditional quantile is then given as

$$\hat{q}^{np}_\alpha(\xi_{\tau+1}|\xi_\tau) = \hat{q}^{np}_\alpha(y|x) = -\hat{F}^{-1}_{np}[\alpha|x], \tag{1.20}$$

[22] For the Student$-t$ distributions the degrees of freedom are re-estimated at each forecast origin.

[23] In this study, the Gaussian kernel defined as $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2)$ and a bandwidth $h = 1.06(W-1)^{-0.2}$ are used throughout and yield satisfactory results. An optimal bandwidth rule for conditional quantile estimation can be found in Cai (2002).

where the approximation is along the grid vector $\boldsymbol{x} = (-4.0, -3.9, \ldots, 4.0)'$ and linear interpolation is applied to determine $\hat{q}_\alpha^{np}(y|x)$, where $y = \xi_{\tau+1}$ and $x = \xi_\tau$. With this quantile at hand, the cES can subsequently be estimated by means of the *plug-in method* (Cai and Wang, 2008) as

$$\widehat{\mathrm{cES}}_\alpha^{np}(y|x) = -\frac{1}{\alpha}\int_{-\infty}^{\hat{q}_\alpha^{np}(y|x)} y\hat{f}_{np}(y|x)dy. \tag{1.21}$$

Since the estimated cCDF is discrete, an approximation of (1.21) is

$$\begin{aligned}\widehat{\mathrm{cES}}_\alpha^{np}(\xi_{\tau+1}|\xi_\tau) &= \widehat{\mathrm{cES}}_\alpha^{np}(y|x) \\ &= -\frac{1}{\alpha}\sum_{y_i \le \hat{q}_\alpha^{np}(y|x)} \left(y_i - \frac{y_i - y_{i-1}}{2}\right)[\hat{F}_{np}(y_i|x) - \hat{F}_{np}(y_{i-1}|x)]. \end{aligned}\tag{1.22}$$

To evaluate conditional quantiles and moments of the copula distributions we apply MC techniques.[24] Simulated random variables are subjected to nonparametric evaluation of the conditional distribution of interest, i.e.,

$$\widehat{F}(y|x) = \frac{1}{S}\sum_{s=1}^{S}\hat{F}^{(s)}(y|x) \text{ with } \widehat{F}^{(s)}(y|x) = \frac{\sum_{t=1}^{T-1} K_h(x - v_{2,t}^{(s)})I(v_{1,t}^{(s)} \le y)}{\sum_{t=1}^{T-1} K_h(x - v_{2,t}^{(s)})}. \tag{1.23}$$

Kernel based estimates $\widehat{F}^{(s)}(y|x)$ are obtained from the same $S = 1{,}000$ simulations employed to determine the discretized distributions $H_{(m,n)}^{(j)}$ in (1.15). Quantiles and tail expectations for simulated data are obtained in full analogy to the empirical counterparts given in (1.20) and (1.22). For further details on these simulations see section 1.3.2.

1.4.3 Model combination

In a broad literature initiated by Bates and Granger (1969) and reviewed, e.g., by Timmermann (2006), *forecast combinations* are seen as a promising means to exploit model complementarities, and to cope with various kinds of (unknown) model misspecification. In consequence, combined predictions have been suggested to im-

[24] Although the model framework is parametric, the analytical extraction of quantiles and moments from the copulae is not straightforward due to the standardization step in transforming correlated marginals to sMDS innovations.

prove upon the forecast performance of single model specifications. Therefore, we further consider three risk predictors that combine the informational content of single forecasts. Firstly, to cope with potential (local) misspecification arising from our matching search, we use at each forecast origin τ average risk predictions obtained from the five standardized copula distributions that yield the smallest $D^{(j)}$ statistics given in (1.13). We refer to this forecast combination as 'Co1'. Secondly, we combine the four risk predictors relying either on IID innovations (Gaussian and standardized Student-t quantiles), or on serial dependence patters ($F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ and $F_{np}(\xi_{\tau+1}|\xi_\tau)$). This combined predictor denoted as 'Co2' is determined such that at each forecast origin τ the particular risk model is chosen that has offered the best empirical cVaR coverage over the most recent 250 time instances. Similarly, conditioning on the best cES predictive accuracy, we combine the two local risk predictors $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ and $F_{np}(\xi_{\tau+1}|\xi_\tau)$. The weighting applied to determine the combined predictor 'Co3' is derived from the relative inverse mean absolute cES forecast error over the most recent 250 time instances.

1.4.4 Forecast evaluation

While the evaluation of cVaR models has seen numerous applications, tests on cES forecast accuracy have been applied less frequently (see McNeil and Frey, 2000, Angelidis and Degiannakis, 2007, Diks et al., 2011 and Zhu and Galbraith, 2011).

Generally, the cVaR backtesting relies on so-called *VaR hits*, i.e.

$$hit^{\bullet}_{\tau+1,\alpha} = I(r_{\tau+1} \leq -\widehat{\mathrm{cVaR}}^{\bullet}_{\alpha}(r_{\tau+1}|\Omega_\tau)),\ \tau = W, \ldots, T-1. \tag{1.24}$$

Since the cES is only defined for $r_{\tau+1} \leq -\widehat{\mathrm{cVaR}}^{\bullet}_{\alpha}(r_{\tau+1}|\Omega_\tau)$, a correct specification of the hit series is required prior to cES forecast evaluation. Thus, cES tests potentially suffer from data scarcity, since they depend on the condition of observing a cVaR hit first.[25] For this reason, we put particular emphasis on the evaluation of cES forecasts by means of both common statistical loss functions, and a more economic criterion that is motivated from the perspective of a representative regulator. Next, we outline a prominent cVaR diagnostic, and turn to alternative tools for cES evaluation.

[25] To deal with this problem Zhu and Galbraith (2011) define a hit as a return below a certain threshold loss of -1.2% to -0.6%. Thereby they differentiate cES from cVaR diagnosis. These threshold levels are, however, not sufficiently conservative for practical applications.

1.4.4.1 Value-at-risk diagnosis

For the cVaR evaluation we first apply the *likelihood ratio* (LR) test by Kupiec (1995) on the correct *unconditional* cVaR coverage of a forecast model. The LR test thus assesses whether the cVaR is violated more or less often than 100α percent of the times. The corresponding test statistic reads as $LR = 2\log\left(\left(\frac{1-\hat{\alpha}}{1-\alpha}\right)^{(1-\hat{\alpha})T}\left(\frac{\hat{\alpha}}{\alpha}\right)^{\hat{\alpha}T}\right)$ where $\hat{\alpha} = \frac{1}{T-W}\sum_{\tau=W}^{T-1} hit^{\bullet}_{\tau+1,\alpha}$ and follows a $\chi^2(1)$ distribution. Second, we also apply the *dynamic quantile (DQ)* test by Engle and Manganelli (2004). The DQ test extends the LR test and formalizes both a correct unconditional and conditional VaR specification under the null hypothesis. In this framework it is tested if the centered hits defined as $\text{hit}^{\bullet}_{\tau+1,\alpha} = hit^{\bullet}_{\tau+1,\alpha} - \alpha$, follow a Martingale Difference Sequence (MDS). We restrict our analysis to the following regression model including five lags of centered hits at the nominal coverage level α and one lag of the hits at the full set of considered nominal coverage levels

$$\text{hit}^{\bullet}_{\tau+1,\alpha} = \beta_0 + \sum_{k=2}^{5}\beta_k \text{hit}^{\bullet}_{\tau-k+1,\alpha} + \theta_1 \text{hit}^{\bullet}_{\tau,.01} + \theta_2 \text{hit}^{\bullet}_{\tau,.025} + \theta_3 \text{hit}^{\bullet}_{\tau,.05} + u_{\tau+1}. \quad (1.25)$$

The null hypothesis of correct conditional and unconditional coverage of the model, hence, reads as $H_0 : \beta_k = 0,\ \theta_l = 0\ \forall\ k = 0, 2, \ldots, 5; l = 1, 2, 3$.[26]

1.4.4.2 Expected shortfall diagnosis

The cES forecast diagnosis relies on two measures. First, similar to Angelidis and Degiannakis (2007) and Zhu and Galbraith (2011), we evaluate the overall cES forecasting accuracy according to the *mean-absolute error* (MAE). In order to allow a comparison of the cES forecasting accuracy across models, it is necessary to define a common sequence of cVaR hits. This is due to the fact that the sequence of cVaR hits is likely to differ across models with more hits for the least conservative forecast and substantially fewer hits for more conservative forecasts. We therefore define a

[26] As a further specification of the DQ regression we have considered a purely autoregressive design for the centered hit process. Diagnostic results do not change qualitatively in such a more restrictive setting.

joint cVaR hit to occur if the loss in $\tau+1$ exceeds all four rival cVaR forecasts, i.e.

$$hit^{*}_{\tau+1,\alpha} = hit^{\Phi}_{\tau+1,\alpha} \times hit^{t(\nu)}_{\tau+1,\alpha} \times hit^{np}_{\tau+1,\alpha} \times hit^{j^{*}}_{\tau+1,\alpha}. \tag{1.26}$$

Thereby, we ensure comparability of the cES forecasting performance by means of absolute forecast errors.[27] The MAE is defined as $MAE^{\bullet}_{\alpha} = \frac{1}{\widetilde{T}} \sum_{\tau=T-W}^{T-1} \left| l^{\bullet}_{\tau+1,\alpha} \right|$, with $\widetilde{T}$ denoting the number of joint cVaR hits and

$$l^{\bullet}_{\tau+1,\alpha} = \begin{cases} (|r_{\tau+1}| - \widehat{\mathrm{cES}}^{\bullet}_{\alpha}(r_{\tau+1}|\Omega_{\tau})) & \text{if } hit^{*}_{\tau+1,\alpha} = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{1.27}$$

Second, as Granger and Pesaran (2000) point out, the choice of a loss functional should reflect the objectives of the forecast user. Therefore, we follow Sarma et al. (2003) and consider a potential loss function of a regulating institution to evaluate the cES forecast performance. Noticing that only negative news put an investor's balance sheets under stress, this loss function focuses at situations where a loss exceeds the predicted cES. Employing the cES as a regulation criterion, a *regulator's loss* (RL) may read as

$$RL^{\bullet}_{\alpha} = \frac{1}{\breve{T}^{\bullet}} \sum_{\tau=T-W}^{T-1} \left| l^{r,\bullet}_{\tau+1,\alpha} \right|, \tag{1.28}$$

with $\breve{T}^{\bullet}$ denoting the model-specific number of cES hits, and

$$l^{r,\bullet}_{\tau+1,\alpha} = \begin{cases} (|r_{\tau+1}| - \widehat{\mathrm{cES}}^{\bullet}_{\alpha}(r_{\tau+1}|\Omega_{\tau})) & \text{if } r_{\tau+1} \leq -\widehat{\mathrm{cES}}^{\bullet}_{\alpha}(r_{\tau+1}|\Omega_{\tau}) \\ 0 & \text{otherwise.} \end{cases} \tag{1.29}$$

In order to test if two models differ significantly in their forecasting accuracy, either with regards to their MAE or to their RL, we employ the finite-sample *sign test* as proposed by Diebold and Mariano (1995). This test assesses the null hypothesis of equal predictive ability of two competing models. For this, the null hypothesis is a zero-median loss differential $H_0 : \mathrm{med}(l^{\cdot,\Phi}_{t,\alpha} - l^{\cdot,\bullet}_{t,\alpha}) = 0$, where $l^{\cdot,\Phi}_{t,\alpha}$ is either the MAE or RL loss of the Gaussian benchmark model. The test statistic is then given by

[27] Instead, considering, e.g., the union set of cVaR hits, one would give an advantage to the least conservative forecast, since most excess losses can be expected to occur close to the cVaR forecast. As this model's cES forecast is also likely closest to the cVaR forecast, it would show the best cES forecasting accuracy by construction.

$S_2 = \sum_{t=\tau-W+1}^{T} I_+(d_{\Phi,\bullet,t})$, where $d_{\Phi,\bullet,t} = l_{t,\alpha}^{;\Phi} - l_{t,\alpha}^{;\bullet}$ and $I_+(d_{\Phi,\bullet,t}) = 1$ if $d_{\Phi,\bullet,t} > 0$. In large samples, the studentized S_{2a} statistic $S_{2a} = \frac{S_2 - 0.5T}{\sqrt{0.25T}}$ is standard normal. The standard Diebold-Mariano (DM) test is not defined here, as the set of cVaR hits is not continuous. For the application of the sign test in our setting, it is again decisive that the sets of hits for both competing cES forecast models under consideration are identical.[28]

1.5 Empirical results

In this section we first describe the performance of particular groups of copula based models (copula families, choice of marginals) in approximating local time windows of TGARCH innovations. In light of the diagnosed dependence patterns, it is then of particular interest if cVaR or cES forecasts can be improved by accounting for serial dependence patterns in TGARCH innovations. Related to the ongoing discussion in the Basel framework, we focus on results for the nominal coverage levels $\alpha = 0.01$ and $\alpha = 0.025$. Nonetheless, we also provide the results for the $\alpha = 0.05$ coverage level with brief discussions in footnotes. Diagnostic results are shown for each market. To facilitate the evaluation of overall performance, we also provide suitable summary statistics. If not stated explicitly, the discussion of the inferential results refers to the five percent significance level.

1.5.1 Copula selection

Complementing the results for the in-sample analysis in Table 1.B.5, Table 1.B.6 documents how often particular combinations of the copula families (Clayton, Frank, Gumbel) and marginal distributions (Gaussian, Student$-t(\nu)$, $\nu = 5, 10, 15, 20$) obtain the smallest distance statistics $D^{(j)}$ given in (1.13) when applied to rolling windows of length $W =$ 1,000.

Similar to the full sample matching, the copula distributions allowing for higher order dependence fit the data better than presuming IID distributions for out-of-sample modeling at each market. Again, the best description of the innovation

[28] The sign-test for significant differences in the MAE criterion is also based on the set of joint hits $hit^*_{\tau+1,\alpha}$. For the RL criterion, we employ the joint set of cES hits defined in (1.29) between each forecast model and the Gaussian cES forecast.

sequences is obtained by a representative of the Clayton copula with standardized Student-$t(10)$ marginal distribution, and the Gumbel family with $t(5)$ or $t(10)$ marginals. At an overall level, some Clayton and Gumbel model specifications are selected with equal frequency in about 45% of all ex ante predictions. The Clayton and Gumbel copula with $t(10)$ marginal distribution account for largest shares of 25.70% and 22.27%, respectively. While copula selection results are generally in favor of the Clayton and Gumbel model, conditioning on particular markets reveals the heterogeneity of serial dependence patterns.

1.5.2 cVaR forecasting

To see if the unconditional coverage of cVaR forecasts derived from $F_{np}(\xi_{\tau+1}|\xi_\tau)$, $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$, or combinations thereof is superior to that of risk forecasts relying on IID innovations, we first report empirical coverage estimates for risk levels $\alpha = 0.01$ and $\alpha = 0.025$ in Table 1.B.7.

Unconditional cVaR coverage levels differ strongly across markets. Violations of the nominal coverage level seem to be the rule rather than the exception, with most coverage levels exceeding the nominal level. This indicates that, in general, forecasts are not sufficiently conservative. We find that the detected misspecification is statistically significant for the majority of markets for the Gaussian model as indicated by the Kupiec (1995) LR test. While accounting for leptokurtosis by means of the IID Student$-t$ distribution reduces the unconditional misspecification markedly (1.98% to 1.54%), further improvements are obtained when exploiting higher order dependence by means of nonparametric estimates $F_{np}(\xi_{\tau+1}|\xi_\tau)$ or copula matchings $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$. This reduces the misspecification sizeably and renders it largely insignificant. These gains are largest for the most conservative coverage level $\alpha = 0.01$, where forecasts from combining the best fitting copula model and the nonparametric estimates (Co3) reduce the misspecification of the Gaussian forecast by almost one half (i.e. the unconditional coverage shrinks from 1.98% to 1.04%). For $\alpha = 0.01$, forecasts relying only on copula models (F_{j*} or Co1) provide the best unconditional coverage for six markets. The best coverage is, however, achieved by combining F_{j*} with F_{np} (Co3) which performs best for nine markets. Overall, there are just two markets (BSE Sensex 30 and JKSE) where Gaussian or Student-t quantiles uniquely perform best. For $\alpha = 0.025$ the gains are smaller in relative

terms with Co3 reducing the misspecification on average by about one fourth. For this coverage level, however, the Gaussian or Student-t forecasts are uniformly outperformed (or matched for the BSE Sensex 30) by at least one forecast that exploits serial dependence patterns.[29] All in all, predictors exploiting the dependence in subsequent innovations or combinations thereof result in improved unconditional coverage for the majority of markets and for each coverage level. This result is of particular importance, since unconditional coverage is the relevant backtesting criterion in the current Basel framework (Basel Committee on Banking Supervision, 2016). For all nominal significance levels, combining $F_{np}(\xi_{\tau+1}|\xi_\tau)$ and $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ (Co3) obtains average coverage levels that are closer to the nominal benchmark than any other cVaR predictor, especially the IID Gaussian or Student-t forecasts.

We summarize the significance levels of the DQ test applied to each cVaR forecast across all markets by means of Fisher's combined probability test (Fisher, 1934).[30] The insignificant Fisher test statistics documented in the bottom rows of Table 1.B.9 show that $t(\nu)$ and $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ quantiles (including combinations Co1 and Co3) offer best suited risk predictions for $\alpha = 0.01$ with regard to the conditional accuracy of their cVaR forecasts. For the medium nominal level $\alpha = 0.025$, however, copula forecasts and their combination (Co1) offer the best conditional nominal coverage while misspecification is significant for the IID Student$-t$ based quantiles. The Fisher test statistics for both combinations that may in parts rely on the Gaussian, Student-t and the nonparametric forecasts (Co2 and Co3) are still insignificant, yet respective p-values are markedly smaller. The Gaussian and the nonparametric cVaR forecasts violate the hypothesis of correct conditional coverage for both nominal levels $\alpha = 0.01$ and $\alpha = 0.025$.[31]

Summarizing the performance of rival cVaR estimates, we note that accounting for dependence of TGARCH innovations shows strong potential to markedly improve

[29] The results for $\alpha = 0.05$ are provided in Table 1.B.8. The main results of improved unconditional coverage from copula or nonparametric forecasts are robust, with the IID Gaussian or Student-t forecasts being superior in only two markets. However, the gains are somewhat smaller and Co2 becomes the second best approach.

[30] The Fisher test statistic X^2_{2N} is given by $X^2_{2N} = -2\sum_{n=1}^{N}\log(p_n)$ and follows a $\chi^2(2N)$ distribution, where $N = 18$ is the number of stock market indices.

[31] Regarding the less conservative level $\alpha = 0.05$, the Fisher test statistic with weakest significance (32.2%) is obtained when combining the best five copula forecasts $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ (Co1) (see Table 1.B.10). In general, conditional coverage is less likely fulfilled but the patterns of significant diagnostics are similar to those for the coverage level $\alpha = 0.025$.

common risk measures. The diagnostic tests, however, indicate some remaining misspecification. As a consequence, the following discussion of cES estimates does not condition on pretest survival models but sheds light on the performance of all alternative approaches to risk assessment.

1.5.3 cES forecasting

We first investigate if cES forecasts are sufficiently conservative. For this purpose, Table 1.B.11 shows the mean cES forecast error for the joint set of cVaR hits for $\alpha = 0.01$ and $\alpha = 0.025$. For all markets, models and coverage levels, the large majority of forecasts is not sufficiently conservative. On average, the occurred losses exceed the predicted cES. Here, conditional Gaussian forecasts provide the least conservative cES forecasts. While all other models also underestimate the cES in general, this problem is considerably alleviated when risk predictors exploit serial dependence patterns of TGARCH innovations. Forecasts derived from $F_{j*}(\xi_{\tau+1}|\xi_\tau)$ generally perform best across all markets for both coverage levels, but in particular for $\alpha = 0.01$. For $\alpha = 0.025$ nonparametric forecasts $F_{np}(\xi_{\tau+1}|\xi_\tau)$ are almost equally conservative.[32]

While we find that the copula forecast model $F_{j*}(\xi_{\tau+1}|\xi_\tau)$ delivers the most conservative cES forecasts on average, it is of further interest whether this also translates into more accurate forecasts. Conditioning on hits that are common to all cVaR prediction models ($hit^*_{\tau+1,\alpha}$ in (1.26)), Table 1.B.13 shows the mean absolute cES forecast errors relative to the Gaussian MAE.

The results underscore that the Gaussian TGARCH model performs poorly and is outperformed by cES predictions that rely on unconditional tail expectations of standardized Student$-t$ distributed innovations. However, for this loss function, combining the informational content of tail expectations of $F_{np}(\xi_{\tau+1}|\xi_\tau)$ and $F_{j^*}(\xi_{\tau+1}|\xi_\tau)$ (Co3) again yields, on average, smallest relative MAE statistics for $\alpha = 0.01$. At the single market level and with respect to a nominal coverage level of $\alpha = 0.01$ ($\alpha = 0.025$), it turns out that loss statistics obtained from Co3 are smaller than loss statistics from the unconditional $t(\nu)$ model for eleven (eight) out of 18 stock markets. Unconditional forecasts of the standardized $t-$distributed innova-

[32] For $\alpha = 0.05$ $F_{np}(\xi_{\tau+1}|\xi_\tau)$ is slightly more conservative than $F_{j*}(\xi_{\tau+1}|\xi_\tau)$, however, the gains for all forecasts exploiting the dependence patterns are substantial compared with IID forecasts (see Table 1.B.12).

tions and combinations thereof (Co2) perform best in only seven and three markets for $\alpha = 0.01$ and $\alpha = 0.025$, respectively. Hence, there is considerable potential to improve cES forecasting accuracy by deviating from the assumption of IID innovations, especially for a nominal coverage of $\alpha = 0.025$ which has been considered to be used as the new standard risk level (Basel Committee on Banking Supervision, 2016).[33]

To assess whether these gains can also be exploited in an economically meaningful way, Table 1.B.15 documents prediction results for the regulator's loss criterion defined in (1.29). Again results are provided for each prediction strategy relative to losses invoked by cES predictions that are derived from the Gaussian model.

Over all markets and coverage levels the Gaussian TGARCH model is again characterized by the weakest performance. Not surprisingly, determining the cES under the assumption of IID standardized Student$-t$ innovations reduces regulator's losses over all stock markets, on average, by 53% and 32% for nominal coverage levels of $\alpha = 0.01$ and $\alpha = 0.025$, respectively. Risk predictors that exploit serial dependence patterns, however, are uniformly superior (except for the SSEC and $\alpha = 0.01$) to cES forecasts determined from tail expectations of standardized Student$-t$ innovations. With regard to market specific quotes of relative regulator's losses, the copula-specific forecasts $F_{j*}(\xi_{t+1}|\xi_t)$ and the combination consisting of the five best copula models (Co1) result in lowest mean regulator loss for ten ($\alpha = 0.01$) and nine ($\alpha = 0.025$) out of 18 stock markets. For an additional four (five) markets, Co3 performs best for $\alpha = 0.01$ ($\alpha = 0.025$).[34]

In summary, we find that exploiting the dependence of subsequent innovations by non-parametric estimation and copula-based methods leads, on average, to more conservative cES forecasts. It turns out that these forecasts reduce absolute losses and losses exceeding the cES forecast, and yield highest predictive accuracy across the entire set of rival risk assessments.

[33]For the coverage level $\alpha = 0.05$, the Gaussian forecasts are, on average, only outperformed by the IID forecasts from the t-distribution (see Table 1.B.14). It is important to keep in mind though, that this analysis does not condition on the cVaR pretest survival, where forecast models assuming independent innovations showed the strongest patterns of misspecification.

[34]For $\alpha = 0.05$, the gains are again smaller, but still pronounced in the range of one third (see Table 1.B.16). The best forecast model is the combination Co3, obtaining the smallest regulator's loss in eight out of 18 markets.

1.6 Conclusion

Innovations of common (T)GARCH models appear to lack independence which is a common assumption in this modeling framework. For a large cross section of stock market indices we illustrate and detect marked higher order dependence of innovations beyond zero correlation. To address if such serial dependence patterns can be exploited for ex ante risk prediction we perform an out-of-sample analysis for a cross section of 18 major stock markets and evaluate about 18,000 one-step ahead risk forecasts. For this, we quantify the distribution of innovations conditional on their past either nonparametrically or by means of a model class comprising standardized copula distributions. We compare risk forecasts for Value-at-Risk and expected shortfall from these semi-parametric modelling frameworks against rival schemes that rely on the conventional assumption of (leptokurtic) independently and identically distributed innovations. Our empirical performance evaluation also covers combined risk predictors.

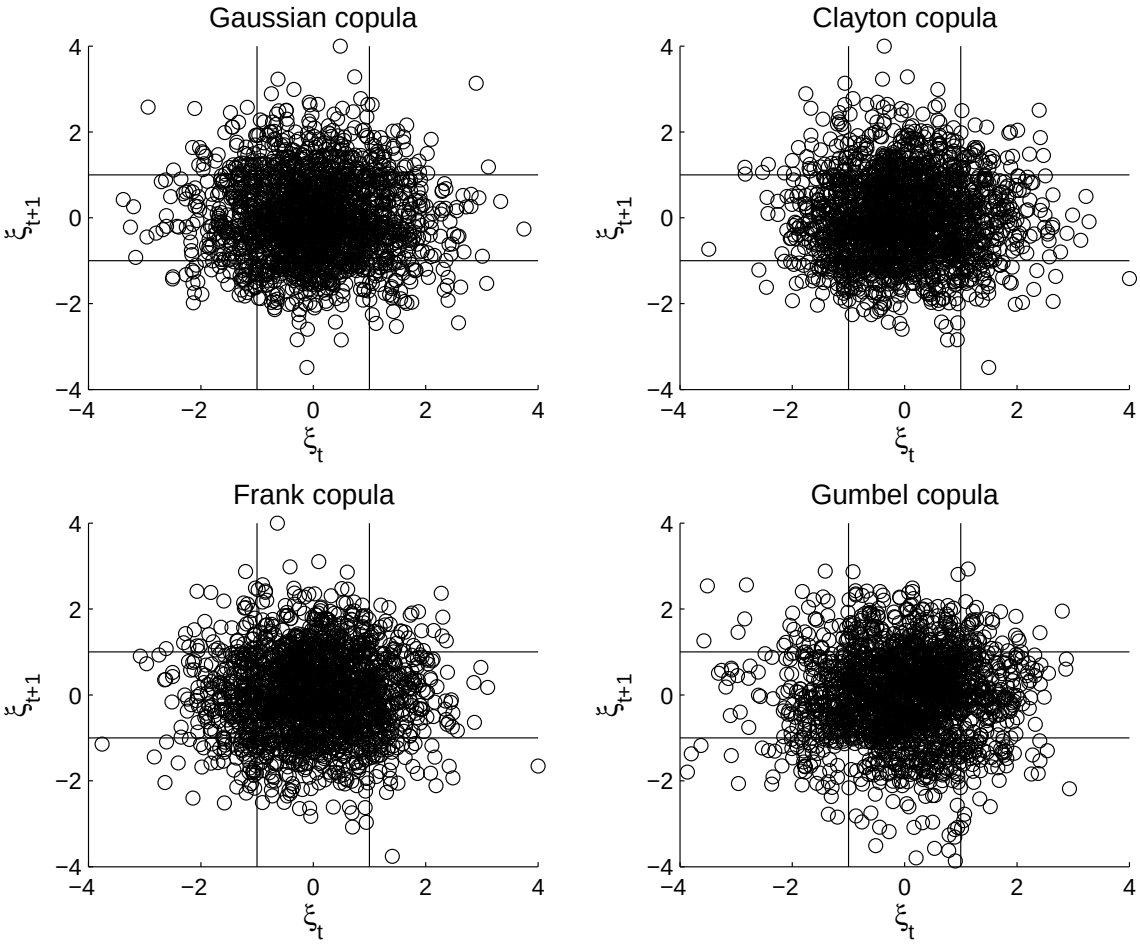
We find that ex ante risk prediction by means of TGARCH models gains from the consideration of potential serial dependence of underlying innovations. Approximating the distribution of innovations conditional on their past by means of flexible copula distributions yields overall more accurate risk forecasts than developing the latter from unconditional quantiles or expectations of standardized Student$-t$ innovations. Among standardized copula distributions the Clayton and Gumbel families, coupled with standardized Student$-t$ marginals outperform the Frank family in terms of in-sample fitting and out-of-sample prediction. In terms of risk prediction, these copula predictors largely outperform their nonparametric counterparts. Combining these two new avenues of risk prediction offers a promising approach to ex ante risk determination.

With regard to future research we notice that the set of applied copula distributions in this work has been limited to the Clayton, Frank and Gumbel families. Candidate marginal distributions have been the Gaussian and the standardized Student$-t$ distribution. Hence, further improvements in risk prediction might be expected from increasing the set of copulae or marginal distributions. Moreover, in light of our empirical analysis it appears promising to develop a powerful criterion to pretest for serial dependence of (T)GARCH innovations to improve risk predictors.

Appendix

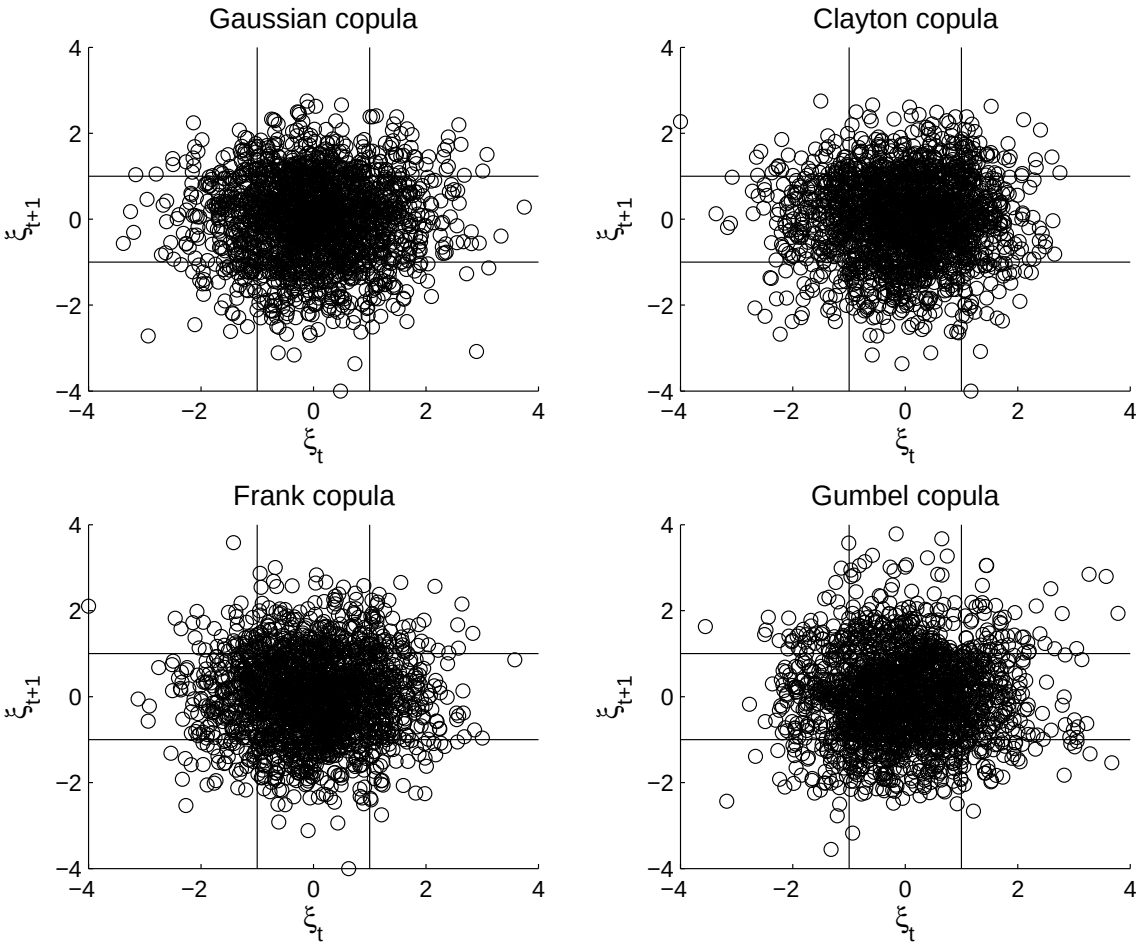
1.A Figures

Figure 1.A.3: Conditional distributions from *unrotated* standardized bivariate Gaussian and Archimedean copulae.



Notes: The figure shows subsequent realizations from a random draw of length $T = 2,000$ from unrotated standardized bivariate Gaussian and Archimedean copulae with Gaussian marginals and a dependence parameter $\tau = 0.250$ prior to standardization.

Figure 1.A.4: Conditional distributions from *rotated* standardized bivariate Gaussian and Archimedean copulae.



Notes: The figure shows subsequent realizations from a random draw of length $T = 2,000$ from rotated standardized bivariate Gaussian and Archimedean copulae with Gaussian marginals and a dependence parameter $\tau = 0.250$ prior to standardization.

1.B Tables

Table 1.B.1: Data set, TGARCH parameter estimates, Ljung and Box (1978) test for serial correlation and Lundbergh and Teräsvirta (2002) test for remaining heteroscedasticity

Index	Country	T	γ_0	γ_1	γ_1^-	β_1	ν	LB(20)	LM(1)	LM(5)
BSE Sensex 30	IN	2228	9.1E-06	.024	.204	.848	9.699	.388	.309	.057
CAC 40	FR	2307	2.4E-06	.000	.162	.906	13.290	1.000	.074	.133
DAX 30	DE	2299	2.5E-06	.000	.149	.914	11.525	1.000	.070	.171
DJIA 30	US	2265	2.0E-06	.000	.178	.894	9.134	.898	.199	.001
Euro Stoxx 50	EZ	2316	2.0E-06	.000	.156	.911	11.270	1.000	.093	.160
FTSE 100	GB	2273	2.0E-06	.000	.169	.899	13.941	.993	.214	.134
IBEX 35	ES	2286	2.0E-06	.000	.152	.914	9.449	1.000	.027	.026
IBRX 50	BR	2214	8.8E-06	.017	.119	.895	10.393	.991	.316	.169
IPC 35	MX	2266	3.8E-06	.012	.149	.895	9.471	.747	.317	.171
JKSE	ID	2201	2.1E-05	.040	.248	.746	6.151	.001	.021	.011
KOSPI	KR	2229	6.9E-06	.000	.175	.877	11.293	1.000	.056	.034
MIB 40	IT	2180	2.0E-06	.000	.163	.905	8.794	1.000	.199	.181
Nasdaq 100	US	2266	3.0E-06	.000	.132	.919	11.904	.992	.160	.183
Nikkei 225	JP	2206	5.1E-06	.021	.139	.883	15.265	1.000	.273	.310
RTSI 50	RU	1873	1.2E-05	.054	.118	.867	5.371	.282	.317	.116
S&P 500	US	2266	2.0E-06	.000	.165	.901	8.884	.799	.217	.004
S&P TSX	CA	2233	2.0E-06	.005	.137	.906	11.385	1.000	.046	.020
SSEC	CN	2273	2.0E-06	.046	.014	.943	4.932	1.000	.114	.198

Notes: Columns γ_0 to β_1 provide the parameter estimates of the TGARCH(1,1,1) model. ν gives the estimated degrees of freedom for the Student$-t$ distribution of innovations. The column labeled 'LB(20)' displays p-values for the Ljung and Box (1978) statistic for testing the null hypothesis of no serial correlation in the model residuals. The columns labeled 'LM(1)' and 'LM(5)' provide p-values for the Lundbergh and Teräsvirta (2002) statistics obtained with one or five lags, respectively, for testing the null hypothesis of no remaining heteroscedasticity in the residuals.

Table 1.B.2: Subsample serial autocorrelation in TGARCH residuals

	Subsample													
Market	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
BSE Sensex 30	0.39	0.07	0.02*	0.04*	0.02*	0.01*	0.00*	0.00*	0.02*	0.06	0.05*	0.06	0.46	
CAC 40	0.55	0.65	0.48	0.59	0.41	0.65	0.65	0.69	0.77	0.82	0.77	0.92	0.82	0.55
DAX 30	0.72	0.94	0.84	0.92	0.76	0.64	0.76	0.82	0.82	0.82	0.85	0.95	0.81	
DJIA 30	0.62	0.85	0.57	0.19	0.06	0.01*	0.06	0.09	0.06	0.16	0.13	0.43	0.54	
Euro Stoxx 50	0.54	0.88	0.87	0.83	0.59	0.74	0.65	0.62	0.37	0.61	0.51	0.86	0.73	0.42
FTSE 100	0.36	0.63	0.44	0.17	0.38	0.63	0.73	0.46	0.44	0.75	0.82	0.93	0.84	
IBEX 35	0.67	0.73	0.83	0.52	0.40	0.46	0.36	0.15	0.35	0.61	0.76	0.95	0.56	
IBRX 50	0.03*	0.11	0.46	0.37	0.27	0.22	0.61	0.25	0.53	0.75	0.78	0.53	0.58	
IPC 35	0.41	0.37	0.25	0.20	0.20	0.26	0.10	0.20	0.26	0.55	0.53	0.59	0.72	
JKSE	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.01*	0.02*	0.02*	0.02*	0.29	0.37	
KOSPI	0.57	0.62	0.60	0.49	0.60	0.75	0.83	0.81	0.79	0.69	0.93	0.96	0.98	
MIB 40	1.00	0.99	0.94	0.73	0.64	0.21	0.27	0.47	0.56	0.37	0.55	0.36		
Nasdaq 100	0.69	0.71	0.79	0.63	0.48	0.31	0.58	0.50	0.48	0.83	0.64	0.80	0.90	
Nikkei 225	0.71	0.87	0.80	0.95	0.65	0.30	0.25	0.08	0.03*	0.16	0.29	0.40	0.69	
RTSI 50	0.01*	0.02*	0.03*	0.17	0.13	0.40	0.09	0.02*	0.01*					
S&P 500	0.78	0.88	0.77	0.34	0.12	0.03*	0.18	0.18	0.11	0.31	0.22	0.48	0.52	
S&P TSX	0.79	0.55	0.97	0.98	0.91	0.78	0.83	0.52	0.74	0.94	0.86	0.77	0.87	
SSEC	0.11	0.03*	0.03*	0.12	0.03*	0.02*	0.01*	0.03*	0.11*	0.06	0.17	0.14	0.30	

Notes: The table displays the p-values of the LB-test for serial correlation with 20 lags in residuals ξ_t by Ljung & Box (1978). The columns indicate the respective subsample, with '1' indicating the first subsample spanning the first 1000 observations, '2' indicating the subsample spanning observations $101, \ldots, 1100$, etc. Values indicated by * are significant at the 5%-level.

Table 1.B.3: Subsample remaining ARCH effects in TGARCH residuals

	Subsample													
Index	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
BSE Sensex 30	0.31	0.18	0.06	0.02*	0.01*	0.20	0.18	0.17	0.17	0.22	0.22	0.28	0.30	
CAC 40	0.32	0.15	0.16	0.31	0.31	0.27	0.31	0.26	0.09	0.08	0.03*	0.02*	0.02*	0.02*
DAX 30	0.29	0.30	0.25	0.25	0.18	0.13	0.19	0.14	0.14	0.13	0.12	0.09	0.03*	
DJIA 30	0.20	0.26	0.32	0.31	0.32	0.32	0.31	0.28	0.28	0.24	0.14	0.21	0.15	
Euro Stoxx 50	0.29	0.19	0.19	0.22	0.27	0.15	0.22	0.19	0.15	0.05	0.02*	0.03*	0.02*	0.02*
FTSE 100	0.27	0.31	0.32	0.30	0.29	0.29	0.32	0.31	0.32	0.28	0.20	0.24	0.17	
IBEX 35	0.23	0.22	0.11	0.13	0.22	0.11	0.16	0.08	0.03*	0.01*	0.03*	0.07	0.12	
IBRX 50	0.31	0.30	0.31	0.29	0.29	0.32	0.30	0.32	0.28	0.31	0.27	0.32	0.25	
IPC 35	0.31	0.32	0.32	0.28	0.32	0.28	0.30	0.24	0.25	0.17	0.31	0.31	0.26	
JKSE	0.09	0.06	0.07	0.09	0.11	0.09	0.12	0.11	0.17	0.19	0.20	0.05*	0.05	
KOSPI	0.18	0.31	0.28	0.29	0.31	0.29	0.12	0.06	0.09	0.05*	0.05	0.00*	0.02*	
MIB 40	0.32	0.31	0.32	0.32	0.09	0.16	0.15	0.09	0.09	0.08	0.08	0.09		
Nasdaq 100	0.30	0.32	0.20	0.15	0.31	0.32	0.32	0.23	0.12	0.08	0.03*	0.05	0.05*	
Nikkei 225	0.01*	0.00*	0.00*	0.01*	0.00*	0.01*	0.01*	0.01*	0.02*	0.03*	0.22	0.20	0.17	
RTSI 50	0.19	0.12	0.11	0.17	0.08	0.14	0.12	0.11	0.14					
S&P 500	0.28	0.31	0.32	0.30	0.32	0.31	0.32	0.28	0.24	0.17	0.05	0.16	0.11	
S&P TSX	0.10	0.04*	0.13	0.27	0.17	0.21	0.31	0.27	0.17	0.16	0.16	0.19	0.04*	
SSEC	0.10	0.20	0.10	0.11	0.08	0.05	0.05	0.15	0.32	0.28	0.30	0.20	0.22	

Notes: The table displays the p-values of the LM-test with one lag for remaining ARCH effects in residuals ξ_t by Lundbergh & Teräsvirta (2002). The columns indicate the subsample with '1' indicating the first subsample spanning the first 1000 observations, '2' indicating the subsample spanning observations $101, \ldots, 1100$, etc. Values indicated by * are significant at the 5%-level.

Table 1.B.4: Higher order co-moments and tests for independence between TGARCH residuals

Index	$\overline{\xi_t^2\xi_{t+1}}$	$\overline{\xi_t\xi_{t+1}^2}$	$\lambda_{iid}^{(3)}$	Φ	$t(10)$	$\widetilde{t(10)}$	$\overline{\xi_t^2\xi_{t+1}^2}$	$\overline{\xi_t^3\xi_{t+1}}$	$\overline{\xi_t\xi_{t+1}^3}$	$\lambda_{iid}^{(4)}$	Φ	$t(10)$	$\widetilde{t(10)}$
BSE Sensex 30	-.072	.000	4.24	.1200	.2107	.0458	.913	.523	.271	42.58	.0000	.0005	.0004
CAC 40	-.066	-.032	3.33	.1883	.2521	.2706	0.816	-.110	-.059	8.18	.0423	.0235	.0003
DAX 30	-.104	-.016	8.67	.0130	.0447	.0939	0.805	.095	.113	9.42	.0242	.0122	.0006
DJIA 30	-.050	.026	3.45	.1780	.3065	.3951	0.772	-.191	-.205	16.83	.0008	.0007	.0002
Euro Stoxx 50	-.075	-.033	4.48	.1061	.1875	.3348	0.821	-.037	.006	6.52	.0888	.0331	.0005
FTSE 100	-.063	-.017	2.95	.2284	.3234	.2499	0.859	-.181	-.065	9.10	.0280	.0870	.0178
IBEX 35	-.064	.002	3.65	.1605	.2852	.2487	0.930	.099	.158	4.83	.1847	.4233	.0065
IBRX 50	-.093	.034	10.30	.0058	.0363	.0397	0.863	.241	.070	13.48	.0037	.0609	.0014
IPC 35	-.102	.011	9.41	.0090	.0360	.2011	0.913	.432	.242	30.19	.0000	.0115	.0030
JKSE	-.153	.044	24.71	.0000	.0002	.0027	0.993	.888	.353	125.30	.0000	.0000	.0000
KOSPI	-.092	.034	9.54	.0085	.0358	.0662	0.849	.177	.017	10.91	.0122	.0455	.0036
MIB 40	-.092	-.029	6.14	.0464	.0993	.1999	0.839	-.041	.162	13.01	.0046	.0305	.0012
NASDAQ 100	-.059	.013	3.57	.1670	.2962	.4425	0.770	-.148	-.144	14.33	.0025	.0015	.0002
NIKKEI 225	-.105	.045	13.36	.0013	.0096	.0397	0.939	.350	.126	21.63	.0001	.0685	.0034
RTSI 50	-.101	-.071	7.21	.0271	.0541	.1156	0.989	.469	.453	34.97	.0000	.0017	.0061
S&P 500	-.047	.047	5.01	.0814	.1830	.2881	0.751	-.182	-.275	23.45	.0000	.0001	.0000
S&P TSX	-.089	-.033	5.87	.0531	.1016	.0930	0.878	.017	.037	3.00	.3907	.2707	.0174
SSEC	-.040	.146	23.60	.0000	.0003	.0124	1.014	.097	-.015	2.73	.4350	.8464	.0794

Notes: The table displays empirical means of third and fourth order co-moments and dependence diagnostics $\lambda_{iid}^{(3)}$ and $\lambda_{iid}^{(4)}$ given in equation (1.4). The columns labeled Φ and $t(10)$ document $p-$values for testing the null hypothesis of no third or fourth order dependence for these diagnostics with MC based covariance estimates obtained from Gaussian and standardized Student$-t(10)$ distributed IID processes, respectively. The columns labeled with $\widetilde{t(10)}$ show $p-$values for diagnostics $\lambda_{iid}^{(3)}$ and $\lambda_{iid}^{(4)}$ determined for trimmed estimated innovation standardized Student$-t(10)$ processes.

Table 1.B.5: Minimum distances between empirical distributions of TGARCH residuals and simulated copulae

	iid		Clayton					Frank					Gumbel				
Index	Φ	$(t\hat{\nu})$	Φ	$t(5)$	$t(10)$	$t(15)$	t(20)	Φ	$t(5)$	$t(10)$	$t(15)$	$t(20)$	Φ	$t(5)$	$t(10)$	$t(15)$	$t(20)$
BSE Sensex 30	.1003	.0823	.0887	.0762	**.0611**	.0650	.0680	.0827	.1149	.0812	.0802	.0806	.0627	.0752	.0628	.0621	.0622
CAC 40	.0893	.0841	.0847	.1053	.0704	**.0689**	.0701	.0847	.1380	.0878	.0833	.0823	.0840	.0965	.0773	.0768	.0774
DAX 30	.1464	.1302	.1389	.1413	**.1123**	.1142	.1163	.1244	.1683	.1294	.1249	.1240	.1194	.1227	.1183	.1183	.1185
DJIA 30	.1673	.1182	.1384	.0995	.0962	.1051	.1114	.1138	.1329	.1147	.1139	.1142	.0962	.0978	**.0945**	.0956	.0956
Euro Stoxx 50	.1088	.0927	.0987	.0976	**.0779**	.0799	.0827	.0910	.1266	.0909	.0891	.0893	.0852	.0927	.0805	.0809	.0813
FTSE 100	.0959	.0797	.0799	.0892	**.0621**	.0626	.0646	.0781	.1222	.0804	.0786	.0782	.0740	.0822	.0646	.0667	.0682
IBEX 35	.1028	.0782	.1028	.0781	.0695	.0749	.0797	.0776	.1052	.0775	.0768	.0768	.0696	.0731	**.0671**	.0681	.0681
IBRX 50	.1066	.0821	.1020	.0831	**.0667**	.0715	.0763	.0830	.1108	.0811	.0824	.0828	.0724	.0762	.0706	.0724	.0723
IPC 35	.1205	.0817	.1025	.0683	.0648	.0722	.0777	.0774	.0980	.0796	.0785	.0778	.0581	.0625	**.0569**	.0572	.0576
JKSE	.1940	.1086	.1485	.1103	.1200	.1344	.1454	.1088	.1080	.1002	.1002	.1011	.1172	**.0915**	.1012	.1101	.1120
KOSPI	.1566	.1337	.1362	.1217	.1046	.1082	.1117	.1299	.1638	.1321	.1301	.1301	.1062	.1164	.1050	**.1041**	.1043
MIB 40	.1550	.1201	.1041	.0867	**.0779**	.0831	.0871	.1202	.1434	.1197	.1196	.1194	.0806	.0868	.0783	.0789	.0793
NASDAQ 100	.1099	.0955	.1010	.1028	**.0828**	.0846	.0870	.0915	.1304	.0949	.0924	.0917	.0860	.0929	.0853	.0862	.0870
NIKKEI 225	.1301	.1206	.1047	.1481	.1060	.1020	**.1011**	.1101	.1719	.1235	.1168	.1141	.1286	.1297	.1154	.1174	.1189
RTSI 50	.2060	.1114	.1401	.1117	.1248	.1428	.1479	.1186	.1118	.1101	.1101	.1109	.1101	**.0997**	.1064	.1074	.1088
S&P 500	.1568	.1281	.1286	.1148	**.1021**	.1062	.1098	.1217	.1502	.1246	.1225	.1220	.1064	.1076	.1047	.1052	.1052
S&P TSX	.1298	.1097	.0967	.0889	.0696	.0730	.0764	.1095	.1448	.1091	.1095	.1098	**.0696**	.0782	.0697	.0698	.0697
SSEC	.1745	.1068	.1751	.1064	.1143	.1280	.1370	.1000	.1039	**.0978**	.0985	.0987	.1370	.1059	.1152	.1289	.1346
Mean	.1361	.1035	.1151	.1017	.0880	.0931	.0972	.1013	.1303	.1019	.1004	.1002	.0924	.0938	**.0874**	.0892	.0900

Notes: The table displays the minimum distances $D^{(j)}$ defined in (1.13) between discretized distributions of empirical TGARCH innovation and simulated copulae processes. The first two columns display the distances to the IID Gaussian (Φ) and to the IID Student-t distribution with estimated degrees of freedom $\hat{\nu}$ ($t(\hat{\nu})$). Among each class of copulae (Clayton, Frank and Gumbel with marginal distributions Φ or $t(\nu), \nu = 5, 10, 15, 20$), the best parametrization as characterized by Kendall's τ is selected. The best fitting distribution is highlighted with bold entries.

Table 1.B.6: Copula selection

	Clayton					Frank				Gumbel				
Index	Φ	$t(5)$	$t(10)$	$t(15)$	t(20)	Φ	$t(10)$	$t(15)$	$t(20)$	Φ	$t(5)$	$t(10)$	$t(15)$	$t(20)$
BSE Sensex 30	0.00	7.57	24.08	10.09	0.00	0.00	5.29	0.00	0.00	0.00	1.79	**39.30**	9.68	2.20
CAC 40	0.08	0.00	19.80	36.62	**43.50**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DAX 30	0.00	0.00	**42.46**	29.77	25.77	1.15	0.00	0.00	0.00	0.00	0.00	0.62	0.23	0.00
DJIA 30	0.00	0.00	15.40	4.50	0.00	0.55	0.00	0.00	0.16	0.00	0.55	**78.83**	0.00	0.00
Euro Stoxx 50	0.00	0.00	**49.20**	22.10	4.02	7.97	0.00	2.05	2.96	0.00	0.00	6.99	3.95	0.76
FTSE 100	0.00	0.00	**60.20**	18.76	0.08	0.00	0.00	0.00	0.00	0.00	0.00	5.10	11.30	4.55
IBEX 35	0.00	0.00	5.67	5.91	2.25	13.29	2.49	1.01	0.00	0.00	**45.61**	23.78	0.00	0.00
IBRX 50	0.00	0.00	**93.58**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.26	0.16	0.00
IPC 35	0.00	9.94	0.00	0.39	0.00	0.32	0.00	0.00	0.00	0.00	0.00	**64.96**	24.39	0.00
JKSE	0.00	1.25	0.00	0.00	0.00	15.39	0.33	2.91	0.92	0.00	**78.45**	0.75	0.00	0.00
KOSPI	0.00	0.00	29.84	2.52	0.00	0.00	0.00	0.00	0.00	3.74	0.00	**51.30**	4.72	7.89
MIB 40	0.00	0.00	4.57	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.34	**51.48**	26.93
NASDAQ 100	10.89	0.87	11.05	8.92	9.94	13.89	0.00	0.00	0.00	0.00	0.00	**43.88**	0.55	0.00
NIKKEI 225	3.40	0.00	0.00	14.17	**62.05**	20.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RTSI 50	0.00	1.72	12.36	0.00	0.00	16.82	0.00	2.97	3.78	0.00	**47.48**	8.70	6.18	0.00
S&P 500	0.00	8.29	34.49	2.68	3.79	5.13	0.00	0.00	0.00	0.00	6.47	**39.15**	0.00	0.00
S&P TSX	0.00	0.00	**59.89**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.57	16.94	19.61
SSEC	0.00	0.08	0.00	0.00	0.00	13.66	1.65	6.04	25.27	0.00	**41.92**	11.38	0.00	0.00
Mean	0.80	1.65	**25.70**	8.73	8.41	6.03	0.54	0.83	1.84	0.21	12.35	22.27	7.20	3.44

Notes: The table displays the frequency (in %) how often each distributional class is selected in the recursive estimation setting. The IID standard normal and t-distribution with estimated degrees of freedom as well as the Frank copula with $t(5)$ margins never obtain smallest distance statistics and were excluded from the Table. The most frequently selected distributional class is highlighted with bold entries.

Table 1.B.7: Unconditional cVaR coverage levels: $\alpha = 0.01$ and $\alpha = 0.025$

	$\alpha = 0.010$							$\alpha = 0.025$						
Index	Φ	$t(\nu)$	F_{np}	F_{j^*}	Co1	Co2	Co3	Φ	$t(\nu)$	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	**.0102**	.0082	.0082	.0072	.0072	.0072	.0041*	**.0184**	.0164	**.0184**	.0133*	.0133*	.0153*	.0133*
CAC 40	.0180*	.0161	.0180*	**.0132**	**.0132**	.0151	**.0132**	.0397*	.0369*	.0331	.0312	**.0303**	.0341	.0322
DAX 30	.0200*	.0162	**.0095**	.0124	.0124	.0114	**.0095**	.0448*	.0429*	.0305	.0324	.0315	.0305	**.0257**
DJIA 30	.0286*	.0158	.0158	**.0118**	**.0118**	.0128	.0128	.0424*	.0404*	.0315	.0374*	.0374*	.0315	**.0305**
Euro Stoxx 50	.0188*	.0159	.0169*	.0131	.0131	.0150	**.0122**	.0432*	.0385*	.0356*	.0366*	.0347	.0385*	**.0338**
FTSE 100	.0235*	.0186*	.0147	**.0117**	.0127	.0137	.0137	.0381*	.0371*	.0342	**.0323**	**.0323**	**.0323**	.0332
IBEX 35	.0193*	**.0116**	**.0116**	.0058	.0048	.0058	.0029*	.0376*	.0357*	.0290	.0280	.0319	.0290	**.0251**
IBRX 50	.0145	.0114	.0124	**.0104**	**.0104**	.0124	.0093	.0342	.0311	**.0259**	.0270	**.0259**	**.0259**	**.0259**
IPC 35	.0197*	.0148	.0148	**.0108**	**.0108**	.0118	.0118	.0354*	.0335	.0285	**.0266**	.0276	.0285	.0285
JKSE	.0210*	**.0116**	.0137	.0074	.0084	.0084	.0084	.0326	.0284	**.0252**	.0242	**.0252**	.0263	**.0252**
KOSPI	.0174*	.0163	.0163	.0112	.0112	.0123	**.0102**	.0368*	.0358*	**.0286**	**.0286**	.0296	.0306	.0296
MIB 40	.0280*	.0258*	.0226*	**.0161**	.0172*	.0183*	**.0161**	.0441*	.0419*	.0409*	**.0355**	**.0355**	.0376*	**.0355**
Nasdaq 100	.0177*	.0167*	.0157	.0138	.0138	.0167*	**.0128**	.0413*	.0404*	.0295	**.0276**	.0285	.0285	.0285
Nikkei 225	.0188*	.0167	**.0105**	.0115	.0126	.0126	**.0105**	.0345	.0345	.0303	.0293	.0293	.0282	**.0272**
RTSI 50	.0144	.0112	.0161	.0112	.0112	.0128	**.0096**	.0321	.0289	**.0257**	**.0257**	**.0257**	.0225	.0193
S&P 500	.0285*	.0207*	.0167*	.0148	.0148	.0157	**.0138**	.0413*	.0413*	.0335	.0354*	.0364*	.0335	**.0325**
S&P TSX	.0214*	.0173*	**.0102**	.0112	.0112	**.0102**	.0081	.0387*	.0366*	**.0264**	.0305	.0305	.0295	.0275
SSEC	.0166	**.0117**	**.0117**	.0068	.0059	.0078	.0078	.0362*	.0323	.0215	.0264	**.0254**	.0264	.0244
Mean	.0198	.0154	.0142	.0111	.0113	.0122	**.0104**	.0373	.0351	.0294	.0293	.0295	.0294	**.0277**
Stand. Dev.	.0049	.0041	.0036	.0028	.0031	.0034	.0034	.0061	.0063	.0052	.0056	.0055	.0053	.0054

Notes: The table displays the empirical cVaR coverage for nominal levels of $\alpha = 0.01$ and $\alpha = 0.025$. $\mathbf{\Phi}$ denotes the IID Gaussian forecast, $t(\nu)$ the IID Student-t forecast with estimated degrees of freedom, F_{np} the forecast with the non-parametric estimate of the conditional distribution of ξ_t, F_{j^*} the forecast using the best-fitting copula distribution, and Co1, Co2, Co3 the forecast combinations described in Section 1.4.3. Bold numbers indicate the most favorable results, where best empirical coverage is defined as the smallest absolute deviation to nominal coverage level. Entries denoted by * deviate significantly from the target coverage level according to the (Kupiec, 1995) test at the 5% level.

Table 1.B.8: Unconditional cVaR coverage levels: $\alpha = 0.05$

Index	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	**.0491**	.0521	.0450	.0409	.0419	.0450	.0440
CAC 40	.0710*	.0728*	**.0643***	**.0643***	.0653*	.0653*	.0662*
DAX 30	.0696*	.0696*	**.0610**	.0677*	.0677*	**.0610**	.0629
DJIA 30	.0650*	.0690*	.0621	.0552	.0552	.0571	**.0542**
Euro Stoxx 50	.0732*	.0732*	.0582	.0675*	.0666*	**.0572**	.0629
FTSE 100	.0635	.0665*	.0557	.0596	.0587	**.0547**	**.0547**
IBEX 35	.0772*	.0782*	**.0743***	.0763*	.0753*	.0772*	**.0743***
IBRX 50	.0560	.0581	.0467	.0529	**.0498**	.0488	.0467
IPC 35	.0541	.0571	.0472	**.0492**	**.0492**	.0472	.0463
JKSE	.0452	**.0505**	.0452	.0431	.0431	.0452	.0442
KOSPI	.0603	.0623	.0511	.0541	.0541	.0562	**.0501**
MIB 40	.0753*	.0796*	.0688*	**.0634**	**.0634**	.0656*	.0645
Nasdaq 100	.0620	.0620	**.0561**	.0600	.0591	.0571	.0581
Nikkei 225	.0638	.0638	.0544	.0554	.0554	.0565	**.0523**
RTSI 50	.0546	.0578	.0546	.0482	.0482	**.0498**	.0514
S&P 500	.0630	.0669*	.0600	.0630	.0620	.0600	**.0591**
S&P TSX	.0722*	.0722*	**.0529**	.0621	.0621	.0549	.0539
SSEC	.0528	.0567	**.0499**	.0528	.0538	.0479	.0489
Mean	.0627	.0649	.0560	.0575	.0573	.0559	**.0553**
Stand. Dev.	.0093	.0085	.0081	.0091	.0089	.0082	.0084

Notes: The table displays the empirical cVaR coverage for a nominal coverage level of $\alpha = 0.05$. See Table 1.B.7 for further notes.

Table 1.B.9: Dynamic Quantile test (p-Values): $\alpha = 0.01$ and $\alpha = 0.025$

	$\alpha = 0.010$							$\alpha = 0.025$						
Index	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	**.998**	.997	.001	.991	.991	.072	.386	**.782**	.589	.104	.182	.181	.053	.181
CAC 40	.475	**.552**	.000	.466	.445	.000	.482	.214	.258	.120	**.435**	.426	.009	.054
DAX 30	.346	.565	.964	.576	.576	.515	**.998**	.046	.086	.163	.611	.672	.524	**.912**
DJIA 30	.024	.602	.606	**.991**	**.991**	.975	.977	.108	.216	.665	.250	.250	**.700**	.660
Euro Stoxx 50	.412	.704	.513	.886	.864	.726	**.980**	.012	.188	.001	.289	**.367**	.061	.257
FTSE 100	.045	.060	.034	.525	**.591**	.476	.494	.278	.134	.253	**.604**	.600	.325	.326
IBEX 35	.313	.518	.001	**.878**	.627	.877	.020	.431	.307	.002	.757	.560	**.831**	.732
IBRX 50	.485	.327	.013	.349	.352	.392	**.999**	.408	**.784**	.074	.617	.516	.023	.011
IPC 35	.060	.013	.005	.481	.481	.000	**.572**	.300	.512	.116	.517	**.538**	.422	.450
JKSE	.360	.994	.655	.994	**.999**	.999	.999	.608	.885	.837	**.912**	.907	.886	.904
KOSPI	.546	.590	.456	.995	.995	.986	**.998**	.178	.228	.310	.762	.702	.718	**.816**
MIB 40	.040	.065	.012	**.848**	.583	.434	.610	.066	.166	.062	.398	.398	.356	**.519**
Nasdaq 100	.392	.399	.564	.948	.949	.602	**.977**	.001	.000	**.829**	.508	.548	.665	.551
Nikkei 225	.394	.592	.357	.494	.588	.584	**.746**	.275	.262	**.619**	.228	.243	.597	.110
RTSI 50	**.259**	.081	.043	.165	.165	.001	.038	.575	**.918**	.395	.902	.902	.352	.774
S&P 500	.021	.309	.410	.637	.639	.622	**.643**	.145	.137	.335	.380	.360	.326	**.403**
S&P TSX	.074	.581	**.998**	.976	.976	.998	.997	.406	.496	**.661**	.523	.523	.393	.273
SSEC	.452	.934	.964	.696	.491	**.996**	.812	.525	**.965**	.906	.957	.959	.665	.923
Fisher-Test	.009	.301	.000	**.999**	.997	.001	.950	.001	.028	.001	**.902**	.895	.109	.331

Notes: The table displays for each market the p-values for the Dynamic Quantile (DQ) test for a correct unconditional and conditional VaR specification for nominal coverage levels $\alpha = 0.01$ and $\alpha = 0.025$ as specified in equation (1.25). The bottom row provides the p-values of the cross-sectional Fisher test (Fisher, 1934) as a summary statistic. Bold entries indicate the most favorable results for each market (largest p-values). See Table 1.B.7 for further notes.

Table 1.B.10: Dynamic Quantile test (p-Values): $\alpha = 0.05$

Index	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	.509	.608	.572	.438	.503	.470	**.665**
CAC 40	.200	.164	.005	**.536**	.523	.049	.165
DAX 30	.148	.149	.274	.207	.205	.604	**.604**
DJIA 30	.422	.261	.269	**.608**	**.608**	.463	.588
Euro Stoxx 50	.227	.232	.274	.232	.270	**.317**	.209
FTSE 100	.423	.197	**.448**	.447	.331	.161	.176
IBEX 35	.066	.065	.016	.090	.094	.063	**.162**
IBRX 50	.776	**.887**	.002	.491	.479	.050	.065
IPC 35	.542	.381	.325	**.544**	.540	.185	.416
JKSE	**.720**	.391	.663	.686	.695	.245	.681
KOSPI	.507	.436	.386	.751	**.758**	.381	.689
MIB 40	.124	.085	.271	.143	.141	.134	**.286**
Nasdaq 100	.025	.025	**.315**	.015	.024	.060	.032
Nikkei 225	**.573**	.566	.510	.425	.440	.333	.221
RTSI 50	.191	.505	.084	.485	.485	.436	**.580**
S&P 500	.233	.208	.323	.255	.240	**.332**	.192
S&P TSX	.043	.043	**.690**	.447	.447	.643	.515
SSEC	.844	.876	.679	.968	.972	**.979**	.804
Fisher-Test	.066	.034	.005	.290	**.322**	.038	.168

Notes: The table displays the p-values for the Dynamic Quantile (DQ) test for a correct unconditional and conditional VaR specification for a nominal coverage level $\alpha = 0.05$ as specified in (1.25), and the p-values of the cross-sectional Fisher test as a summary statistic. See Table 1.B.9 for further notes.

Table 1.B.11: Mean errors for cES forecast: $\alpha = 0.01$ and $\alpha = 0.025$

	$\alpha = 0.010$							$\alpha = 0.025$						
Index	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	-1.343	-.788	-.965	**-.726**	-.735	-.816	-.819	-.859	-.545	-.443	-.450	-.457	**-.421**	-.460
CAC 40	-.505	-.143	-.106	-.063	**-.041**	-.055	-.082	-.380	-.182	-.115	-.077	-.079	**-.065**	-.102
DAX 30	-.653	-.265	-.147	-.177	-.177	-.141	**-.062**	-.558	-.315	**-.139**	-.186	-.190	-.206	-.167
DJIA 30	-.542	**-.026**	.027	.051	.037	.049	.032	-.443	-.163	**.003**	-.047	-.049	-.066	-.024
Euro Stoxx 50	-.443	**.009**	-.092	.083	.073	.094	.016	-.367	-.127	**-.048**	-.071	-.069	-.094	-.053
FTSE 100	-.577	-.262	-.106	**-.019**	-.026	-.049	-.053	-.383	-.218	**-.022**	-.038	-.046	-.084	-.031
IBEX 35	-.408	**.149**	.153	.410	.348	.402	.224	-.352	**-.033**	.207	.230	.181	.182	.222
IBRX 50	-.995	-.486	-.451	-.195	-.211	-.299	**-.156**	-.750	-.446	**-.033**	-.186	-.178	-.069	-.097
IPC 35	-.677	-.193	-.088	.092	.090	.075	**-.004**	-.511	-.252	-.191	.006	**.000**	-.021	-.087
JKSE	-2.458	-1.719	**-1.205**	-1.538	-1.609	-1.511	-1.309	-1.183	-.783	**-.540**	-.586	-.643	-.640	-.566
KOSPI	-1.181	-.761	-.652	**-.407**	-.428	-.426	-.481	-.892	-.634	-.478	**-.283**	-.295	-.419	-.400
MIB 40	-.691	-.202	-.133	.057	.035	.039	**-.035**	-.678	-.378	**-.064**	-.078	-.094	-.082	-.071
Nasdaq 100	-.800	**-.410**	-.520	-.333	**-.319**	-.375	-.436	-.535	-.294	-.302	-.213	**-.209**	-.247	-.258
Nikkei 225	-1.184	-.915	-1.244	-.621	-.579	**-.320**	-.594	-.597	-.454	-.325	-.276	-.264	**-.087**	-.301
RTSI 50	-1.077	-.381	**.001**	-.574	-.582	-.482	-.404	-.813	-.433	**-.138**	-.375	-.380	-.152	-.319
S&P 500	-.613	-.019	-.150	-.031	**.008**	-.046	-.098	-.579	-.262	**-.092**	-.179	-.151	-.096	-.137
S&P TSX	-.410	**.017**	.278	.299	.291	.250	.255	-.434	-.175	.223	.138	**.123**	.176	.179
SSEC	-1.306	**-.198**	-.789	-.365	-.307	**-.452**	-.523	-.654	**.015**	-.319	-.020	-.023	-.255	-.173
Mean	-.881	-.366	-.344	**-.225**	-.230	-.226	-.252	-.609	-.316	-.156	-.150	-.157	**-.147**	-.158
Stand. Dev.	.503	.449	.455	.451	.456	.437	.393	.223	.207	.213	.200	.201	.200	.202

Notes: The table displays the mean cES forecast errors for nominal cVaR coverage levels of $\alpha = 0.01$ and $\alpha = 0.025$. cES forecast errors are determined for the joint set of cVaR hits. Negative values indicate an overly optimistic cES forecast, i.e. that losses on average exceed the cES forecast. Bold numbers indicate the most favorable results defined as the smallest mean absolute errors. See Table 1.B.7 for further notes.

Table 1.B.12: Mean errors for cES forecast: $\alpha = 0.05$

Index	Φ	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	-.246	-.074	.068	.071	.065	**.050**	.071
CAC 40	-.285	-.180	-.068	**-.062**	-.070	-.069	-.064
DAX 30	-.295	-.159	.049	-.022	-.021	-.038	**.015**
DJIA 30	-.375	-.211	**.015**	-.049	-.047	-.050	-.019
Euro Stoxx 50	-.330	-.199	**-.105**	-.128	-.123	-.149	-.116
FTSE 100	-.345	-.251	**-.090**	-.096	-.096	-.109	-.093
IBEX 35	-.139	**.036**	.258	.243	.203	.105	.251
IBRX 50	-.598	-.454	**-.201**	-.272	-.252	-.214	-.234
IPC 35	-.437	-.305	-.165	**-.073**	-.082	-.159	-.118
JKSE	-.765	-.546	**-.310**	-.320	-.382	-.343	-.315
KOSPI	-.647	-.513	-.359	**-.241**	-.248	-.360	-.300
MIB 40	-.533	-.373	**-.099**	-.107	-.109	-.138	-.103
Nasdaq 100	-.406	-.274	**-.099**	-.164	-.160	-.121	-.132
Nikkei 225	-.438	-.355	-.178	-.214	-.212	**-.084**	-.200
RTSI 50	-.330	-.106	.149	.088	**.057**	.121	.119
S&P 500	-.466	-.282	-.057	-.137	-.127	**-.048**	-.099
S&P TSX	-.398	-.277	.123	**-.025**	-.030	.072	.047
SSEC	-.433	-.087	**-.025**	-.074	-.082	-.195	-.053
Mean	-.415	-.256	**-.061**	-.088	-.095	-.096	-.075
Stand. Dev.	.151	.155	.157	.136	.133	.135	.144

Notes: The table displays the mean cES forecast errors for nominal cVaR coverage levels of $\alpha = 0.05$. See Table 1.B.11 for further notes.

Table 1.B.13: Relative MAE for cES forecast: $\alpha = 0.01$ and $\alpha = 0.025$

	$\alpha = 0.010$						$\alpha = 0.025$					
Index	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	.698	.795*	.611	**.608**	.626	.621	.880	1.112*	.885	.878	**.853**	.977*
CAC 40	.924	1.555*	.916	**.908**	1.073	1.117	.938	1.062*	.946	**.920**	.945	.974*
DAX 30	.576	.603	.551	.571	.520	**.287**	.802*	1.014*	.808*	**.789***	.866*	.846*
DJIA 30	.383*	.794*	.367*	.350*	**.337***	.447*	.726	.840*	**.706**	.725	.862	.733
Euro Stoxx 50	.986	1.457*	.970	**.965**	1.002*	1.010*	.932	1.141*	**.879**	.897	1.089*	.996*
FTSE 100	.769*	.804*	.714*	**.654***	.717*	.766*	.822	.916	.805	**.794**	.804	.804
IBEX 35	**.480***	.649*	1.005*	.857*	1.006*	.550*	**.769**	1.117*	.972*	.943*	.984*	.972*
IBRX 50	.698	.647	.540	.567	.615	**.393**	.961*	1.044*	.951*	.956*	1.027*	**.933***
IPC 35	.481	.658	.444	.466	**.425**	.497	.847	.946	.891	.878	.895	**.846**
JKSE	.732	.698*	.726	.733	.715	**.642***	.858	.869*	.863	**.850**	.856	.859
KOSPI	.678	.657	.486	.488	**.479**	.529	.864	.871	.833	.842	.863	**.828**
MIB 40	.661*	.699*	.705*	.706*	.679*	**.632***	.777	.901	.735	**.727**	.762	.790
Nasdaq 100	.740	.758*	.768	.740	.705	**.688**	.942	.873*	.878*	.873*	.884*	**.851**
Nikkei 225	.846	1.252*	.733	.710	**.525**	.681*	.955	1.321*	.869	**.864**	.959	1.075*
RTSI 50	.699	.574*	.805	.732	.707	**.631**	.965	**.806***	1.104	1.099	.814	.884
S&P 500	.597*	.786*	**.514***	.557*	.589*	.580*	.694	.758*	**.662**	.665	.683	.669
S&P TSX	**.625***	.980*	.911*	.889*	.925*	.705*	**.778**	.991	.816	.799	.900	.885
SSEC	**.499**	.947*	.693*	.677	.593	.676*	.984*	1.218*	.856	**.855**	1.148*	.915*
Mean	.671	.851	.692	.676	.680	**.636**	.861	.989	.859	**.853**	.900	.880
Stand. Dev.	.157	.288	.186	.164	.206	.197	.090	.152	.102	.099	.113	.100

Notes: The table displays the MAE for cES forecasts from all candidate models relative to the MAE from the IID Gaussian model for cVaR coverage levels $\alpha = 0.01$ and $\alpha = 0.025$. Values less than one indicate that the candidate model provides more precise cES forecasts than the Gaussian model. Entries denoted by * are significantly different from Gaussian cES forecast at the 5% level according to the sign test. See Table 1.B.7 for further notes.

Table 1.B.14: Relative MAE for cES forecast: $\alpha = 0.05$

Index	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	1.064*	1.212*	1.136*	1.127*	1.103*	1.158*
CAC 40	1.018	1.230*	1.105*	1.084*	1.089*	1.132*
DAX 30	1.025	1.207*	1.101*	1.106*	1.169*	1.141*
DJIA 30	.924	.913*	.935	.942	.943	**.888**
Euro Stoxx 50	.969	1.032*	.970	.975	1.022*	**.954***
FTSE 100	.963	.952*	.943	.948	.982	**.927**
IBEX 35	1.051*	1.196*	1.122*	1.110*	1.118*	1.149*
IBRX 50	.964	1.009*	.954	.957	.973*	**.939**
IPC 35	.917	.961	.904	**.885**	.960	.919
JKSE	**.982**	1.005	1.022	1.002	.999	1.007
KOSPI	.941	.987*	**.914**	.929	.947*	.931
MIB 40	.951	.986*	.929	.937	.945	**.920**
Nasdaq 100	**.991**	1.117*	1.026	1.033	1.073*	1.033*
Nikkei 225	.980	1.19*	.937	**.930**	1.034*	1.037
RTSI 50	1.068*	1.231*	1.241	1.201	1.240	1.194
S&P 500	.911	.829	.914	.900	.811	**.800**
S&P TSX	.923	.999*	**.854***	.863*	.923*	.898*
SSEC	1.005	1.261*	.947	**.937**	.976*	1.064
Mean	**.980**	1.073	.997	.993	1.017	1.005
Stand. Dev.	.050	.132	.103	.096	.101	.114

Notes: The table displays the MAE for cES forecasts from all candidate models relative to the MAE from the IID Gaussian model for a cVaR coverage level $\alpha = 0.05$. See Table 1.B.13 for further notes.

Table 1.B.15: Relative regulator's losses: $\alpha = 0.01$ and $\alpha = 0.025$

	$\alpha = 0.010$						$\alpha = 0.025$					
Index	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	.643*	.775	**.576***	.578*	.634*	.615	.763*	.868	.712*	.712*	**.680***	.766
CAC 40	.621*	1.122	.542*	**.518***	.804*	.667*	.759*	.877*	.655*	**.642***	.738*	.698*
DAX 30	.427*	.616*	.350*	.359*	.339*	**.163***	.689*	.649*	.581*	**.575***	.629*	.585*
DJIA 30	.205*	.437	.130*	.135*	**.117***	.185*	.541*	.413*	.402*	.414*	.502*	**.391***
Euro Stoxx 50	.495*	1.092	**.404***	.413*	.409*	.499*	.664*	.789*	**.564***	.572*	.707*	.606*
FTSE 100	.579*	.503*	.357*	**.334***	.383*	.410*	.703*	.534*	.474*	.478*	.529*	**.465***
IBEX 35	.043*	.690	**.000***	.002*	.008*	**.000**	.429*	.484*	**.208***	.253*	.274*	.219*
IBRX 50	.532*	.569*	.290*	.312*	.491*	**.212***	.799*	.620	.640*	.637*	.632*	**.577***
IPC 35	.372*	.600*	**.149***	.162*	.152*	.244*	.675*	.668*	**.450***	**.450***	.479*	.517*
JKSE	.695*	.585*	.657*	.674*	.646*	**.571***	.760*	.679*	.680*	.697*	.715*	**.669***
KOSPI	.660*	.653*	**.415***	.424*	.419*	.467*	.790*	.723*	**.587***	.598*	.675*	.646*
MIB 40	.460*	.641*	**.300***	.316*	.346*	.329*	.670*	.636*	**.431***	.438*	.522*	.454*
Nasdaq 100	.591*	.665	.559*	**.538***	.554*	.582*	.746*	.720*	.648*	**.643***	.678*	.668*
Nikkei 225	.791*	1.119	.611*	.583*	**.387***	.575	.870*	.979*	.690*	.678*	**.600***	.824*
RTSI 50	.509*	**.361**	.664*	.620*	.583*	.486*	.756*	.570*	.796*	.796*	**.536***	.648*
S&P 500	.288*	.508	.258*	**.249***	.314*	.339*	.556*	.475*	.472*	.450*	**.417***	.441*
S&P TSX	.269*	.139*	.084*	.083*	.145*	**.039***	.590*	.267*	.269*	.276*	.271*	**.260***
SSEC	**.312***	1.016	.466*	.437*	.450*	.524*	.521*	.881	**.476***	.478*	.798*	.615*
Mean	.472	.672	.379	**.374**	.399	.384	.682	.657	**.541**	.544	.577	.558
Stand. Dev.	.194	.269	.202	.194	.206	.204	.115	.183	.156	.149	.150	.162

Notes: The table displays the regulator's loss for cES forecasts from all candidate models relative to the loss from the IID Gaussian model for cVaR coverage levels $\alpha = 0.01$ and $\alpha = 0.025$. See Table 1.B.13 for further notes.

Table 1.B.16: Relative regulator's losses: $\alpha = 0.05$

Index	t(ν)	F_{np}	F_{j^*}	Co1	Co2	Co3
BSE Sensex 30	.846*	.831	**.728***	**.728***	.744	.744*
CAC 40	.870*	.922*	.763*	**.762***	.822*	.783*
DAX 30	.837*	.682*	.705*	.707*	.772*	**.681***
DJIA 30	.770*	**.511***	.593*	.595*	.599*	.532*
Euro Stoxx 50	.823*	.811*	.738*	.735*	.859*	**.714***
FTSE 100	.869*	.684*	.675*	.678*	.714*	**.662***
IBEX 35	.751*	.576*	**.462***	.520*	.692*	.471*
IBRX 50	.875*	.715*	.736*	.724*	.705*	**.700***
IPC 35	.826*	.717*	.596*	**.594***	.711*	.648*
JKSE	.863*	.759*	.753*	.777*	.753*	**.741***
KOSPI	.873*	.788*	**.664***	.678*	.767*	.715*
MIB 40	.848*	.696*	.630*	.636*	.664*	**.622***
Nasdaq 100	.857*	.755*	.763*	.763*	.753*	**.733***
Nikkei 225	.915*	.887*	.764*	.758*	**.709***	.812*
RTSI 50	.840*	.761*	.781*	.782*	.750*	**.719***
S&P 500	.778*	.525*	.644*	.627*	**.506***	.544*
S&P TSX	.824*	**.428***	.509*	.520*	.441*	.455*
SSEC	.710*	.820	**.662***	.663*	.783	.719*
Mean	.832	.715	.676	.680	.708	**.667**
Stand. Dev.	.051	.133	.091	.083	.103	.103

Notes: The table displays the regulator's loss for cES forecasts from all candidate models relative to the loss from the IID Gaussian model for a cVaR coverage level $\alpha = 0.05$. See Table 1.B.13 for further notes.

CHAPTER 2

Detecting Asset Price Bubbles in Real-Time through Indicator Combinations[1]

2.1 Introduction

The U.S. housing market crash resulting in the global financial crisis (GFC) of 2007-08 highlighted the financial stability risks that may emerge from extended periods of excessively high asset price growth. Consequently, this episode has raised the interest of market participants in the early detection of emerging bubbles on asset markets. Likewise, the experience from the GFC has also revived the debate among policymakers and economists about how to contain excessive developments on asset markets in the future and thereby to enhance the stability of the financial system. An important practical prerequisite for employing countercyclical financial market policies is, however, to be able to correctly assess that an asset price appreciation is indeed *excessive*.[2]

[1]I am thankful to Kerstin Bernoth, Jörg Breitung, Uwe Hassler, Helmut Herwartz, Helmut Lütkepohl and Christian Proaño for their helpful comments and advice. Further, I thank participants at the *35th International Symposium on Forecasting 2015*, Riverside; the *IAAE Annual Meeting 2015*, Thessaloniki; the Warsaw International Economic Meeting (WIEM) 2015, Warsaw; the *2015 Annual Conference of the Verein für Socialpolitik*, Münster; the *47th Money, Macro and Finance Research Group Annual Conference*, Cardiff; and the *Workshop "Empirical Macroeconomics"* at Freie Universität Berlin for helpful comments and suggestions on an earlier version of this paper.

[2]Such policies include, for instance, a tightening of the micro- and macroprudential regulatory framework, with the Basel III framework suggesting the use of countercyclical capital buffers, or the debate about whether monetary policy could actively lean against the wind of excessive asset price growth by raising policy rates (see Chapter 4 for a discussion of this debate).

This task of detecting asset price bubbles has long been deemed impossible, especially when to be performed in *real-time* (Trichet, 2005 and Kohn, 2006). For long, the literature has thus identified bubble periods as pronounced price deviations from an (HP-)filtered trend (cf. Detken and Smets, 2004; Adalid and Detken, 2007; Assenmacher-Wesche and Gerlach, 2010). Yet, in a real-time monitoring context, this approach suffers not only from the lack of guidance about the appropriate parametrization of the filter, but also from its commonly known endpoint problem. More recently, however, Phillips et al. (2011) have proposed a promising method that can provide early warning signals for excessive asset price developments, challenging the view that the real-time detection of asset price bubbles is infeasible. Specifically, the authors propose to recursively test both an asset's price and its fundamental series for explosive roots, and show that this real-time monitoring approach is capable of detecting periods that display patterns typical for asset price bubbles. Building on this work, Phillips et al. (2015) have subsequently generalized this monitoring procedure, thereby enhancing its ability to detect the occurrence of multiple, periodically collapsing bubbles within a sample.

This paper contributes to the literature on the detection of asset price bubbles by, first, evaluating this set of popular indicators with regard to their ability to signal bubble emergence and collapse dates in real-time, and, second, by developing two combination approaches to aggregate information across indicators. Applying the individual indicators to U.S. stock market data, I highlight the considerable heterogeneity in the provided signals across all indicators. These results indicate common weaknesses of all individual indicators, especially with regard to their stability during the run of a bubble, their ability to detect its emergence and collapse, and false positive signals during normal times. Moreover, I illustrate that the exact specification of the Phillips et al. (2011, 2015) indicators is crucial. In particular, it matters whether individual price and dividend series or the price-to-dividend ratio are tested for explosive roots – an issue that is not discussed carefully in the literature. These findings present a challenge to policymakers and applied researchers, as the choice of the indicator and its exact specification will affect the received signal.

Despite the heterogeneity in their signals, the run-up to the 1987 stock market crash and the dot-com bubble emerge as common bubble periods signaled by most indicators. For this reason, I propose to combine the individual indicators by ag-

gregating their signals at each margin in the sample in real-time, thus extending the work of Harvey et al. (2015). To outline the benefit of such combinations, I first propose a simple counting approach that can be easily implemented and requires little computational effort. Here, I show that a combination requiring three of in total eight individual indicators to simultaneously signal a bubble detects the key bubble episodes, but avoids issuing false positive signals. Yet, this approach is sensitive to the ad-hoc choice of the threshold and the set of the included individual indicators. Therefore, I subsequently describe how to statistically combine the indicators of Phillips et al. (2011) and Phillips et al. (2015). I implement this by means of the multiple testing procedure by Romano and Wolf (2005), which allows to control for the overall size of such a joint test while also taking the correlation between the individual indicators into account. Applying these combination indicators to U.S. stock market data, I show that they provide a promising approach to accurately detect the emergence and collapse of asset price bubbles, and to improve signal stability during the run of dot-com bubble.

I confirm these findings in a simulation study for bubbly processes with varying length and bubble characteristics. In particular, I first highlight the considerable heterogeneity of bubble signals provided by the individual indicators. Second, I show that all individual indicators are sensitive to the bubble processes' characteristics and feature complementary strengths depending on the timing or frequency of bubbles in the sample. For example, indicators based on price deviations from an HP-filtered trend frequently provide false positive signals, but issue the most stable signal during a true bubble's run. In contrast to this, the Phillips et al. (2011, 2015) indicators issue fewer false signals, but frequently suffer from signal instability. Further, depending on the number of bubbles in the sample and their exact timing, either the Phillips et al. (2011) or the Phillips et al. (2015) indicator are more likely to provide an accurate signal. Finally, both indicators' abilities to detect asset price bubbles depend on whether individual price and dividend series or the price-to-dividend ratio are assessed. The proposed combination indicators make use of these complementary strengths, and by this provide more reliable real-time signals for the emergence and collapse of asset price bubbles, independent of the specific characteristics of the bubble process.

Finally, I show that real-time information about the presence of asset price bubbles can be useful for policymakers and applied researchers – if the provided signal is accurate and reliable. In particular, I here show that the best combination indicators as well as the best specifications for the Phillips et al. (2011, 2015) indicators carry significant additional predictive content for output growth against a benchmark model including other popular real and financial predictors.[3] In contrast, all other indicators generally do not improve the forecasts from the benchmark, emphasizing that the accurate dating of bubble emergence and collapse dates are crucial to exploit the information content from such indicators for forecasting or policymaking. By this, I extend the work Assenmacher-Wesche and Gerlach (2010) who consider the indicator based on the HP-filter to forecast output and inflation. As I show, this indicator generally fails to identify the collapse of an asset price bubble in real-time and frequently provides false positive signals, limiting its use for forecasting.

The remainder of this paper is structured as follows. Section 2.2 introduces the bubble indicator and discusses their specifications. I further illustrate their real-time performance when applied to U.S. stock market data, motivating the idea for indicator combinations. These combination indicators are developed and applied in section 2.3. Subsequently, I evaluate the finite sample power properties of all individual indicators and their combinations in different bubble environments in section 2.4. Finally, section 2.5 discusses the value of these indicators for forecasting output, emphasizing the practical relevance of real-time indicator accuracy. I conclude in section 2.6.

2.2 Detecting asset price bubbles in real-time

I begin by outlining the general framework of asset price bubbles to discuss the basis and main implications for the *real-time monitoring* for excessive asset price developments. An asset price bubble is commonly defined as an upward deviation of the asset's price from its fundamental value (FV) resulting from unjustified beliefs about the asset's price in the future. The definition of a (rational) bubble follows

[3]Further evidence for the real-time predictive content of these asset price bubble is provided in Chapter 3.

from the asset pricing equation

$$P_t = \frac{1}{1+r} E_t \left[P_{t+1} + D_{t+1} \right] \tag{2.1}$$

where P_t, D_t are the asset's market price in and dividend accrued over period t, respectively, and r is the (time-constant) discount rate. Solving the equation forward allows to decompose the current asset price into a fundamental component, $FV_t = \sum_{i=1}^{\infty} \left(\frac{1}{1+r}\right)^i E_t \left[D_{t+i} \right]$, and a bubble component B_t:

$$\begin{aligned} P_t &= \sum_{i=1}^{\infty} \left(\frac{1}{1+r} \right)^i E_t \left[D_{t+i} \right] + \lim_{i \to \infty} \left(\frac{1}{1+r} \right)^i E_t \left[P_{t+i} \right] \\ &= FV_t + B_t, \end{aligned} \tag{2.2}$$

where $B_t > 0$, if the usual transversality condition $\lim_{i \to \infty} \left(\frac{1}{1+r}\right)^i E_t \left[P_{t+i} \right] \to 0$ does not hold.[4] Importantly, for the asset pricing equation given in (2.1) to hold, and for the bubble not to shrink to zero in present value or to outgrow the economy, the bubble component must grow exponentially at the rate $1 + r$, i.e. $B_t = \frac{1}{1+r} E_t [B_{t+1}]$.

This characteristic of exponential growth of the bubble process motivated different testing strategies to detect a bubble. Under the standard assumption that dividends follow a random walk with drift and a bubble is present, i.e. $B_t > 0$, the price series contains an explosive root that results from the bubble process. In contrast, when no bubble is present, i.e. $B_t = 0$, then the price series follows a random walk as does the dividend series. This observation motivated the work of Diba and Grossman (1988) who suggest to exploit this relationship between prices and dividends, and to test for the existence of bubbles using right-sided unit root tests on both series over the full sample.[5] If prices are found to contain an explosive root but dividends follow a random walk, a (rational) bubble must be present. If dividends grow exponentially, too, no inference on the existence of a bubble component is possible.

[4] See Gürkaynak (2008) for a derivation of the model's underlying assumptions. Camerer (1989) and Stiglitz (1990) provide surveys on the theoretical work on the conditions under which rational bubbles can exist.

[5] Other approaches include variance-bound and two-step specification tests for which surveys are provided by Flood and Hodrick (1990) and Gürkaynak (2008). Explicit testing for the existence of a bubble is, however, conceptually limited in general because every test relies on a correct specification of the FV (Flood and Hodrick, 1990). This implies that empirical tests always evaluate the joint null hypothesis of the absence of bubbles and a correctly specified model.

The success of this approach to identify asset price bubbles is, however, limited by its ability to detect periodically collapsing bubbles. This form of a bubble can be nested in the above framework such that the bubble is expected to burst with probability p, or, conditionally on not bursting, to grow at the higher rate $\frac{(1+r)}{1-p}$. Despite such periods of explosive behavior, Evans (1991) shows that the price series as a whole is still integrated of order one such that the approach of Diba and Grossman (1988) fails to detect the bubbly process. This limits the ability of this test to identify bubbles over an entire historical dataset. For policy makers, however, the value of the Diba and Grossman (1988) test is even more limited, as it does not allow to monitor asset prices for the emergence and collapse of a periodically collapsing bubble in *real-time*. Nonetheless, the illustrated framework is helpful for defining the characteristics associated with a (rational) bubble and for providing a structure to think about asset price booms that may raise financial stability risks, i.e., as periods of exponential price growth decoupling from the growth rate of dividend payouts, with a potentially sudden collapse. Therefore, repeated testing for explosive growth in prices and dividends may provide early warning signals for such bubbly periods.

2.2.1 Recursive tests for detecting explosive processes

Addressing the limits of the right-sided unit root test of Diba and Grossman (1988) and the demand for early warning signals about emerging bubbles, Phillips et al. (2011) develop a recursive approach that allows to detect periodically collapsing bubbles and to monitor the emergence and collapse of bubbles in real-time. Specifically, this approach builds on the idea to examine prices and dividends for explosive behavior *at each margin* in the sample by right-tailed unit root tests. More recently, Phillips et al. (2015) generalize this approach with the aim to allow for a flexible "restarting" of the test after a first bubble in the sample has collapsed.[6]

[6] Along the same lines, Homm and Breitung (2012) propose two tests based on structural breaks in the autoregressive parameters or on forecast breakdowns, two of which can also be employed for real-time monitoring and date-stamping. I focus on the two tests by Phillips et al. (2011, 2015) here since Phillips et al. (2015) show that these provide superior signals to the Homm and Breitung (2012) approaches.

The common starting point for the tests of Phillips et al. (2011) and Phillips et al. (2015) is hence a variant of the autoregressive specification

$$z_t = \mu_z + \delta z_{t-1} + \sum_{j=1}^{J} \phi_j \Delta z_{t-j} + v_t, \quad t = 1, \ldots, \tau, \quad v_t \overset{iid}{\sim} N(0, \sigma_v^2) \tag{2.3}$$

where the null hypothesis $H_0 : \delta = 1$ is assessed against its alternative $H_1 : \delta > 1$ using right-tailed ADF tests either on both price and dividend series individually, i.e. $z_t \in \{p_t, d_t\}$, or on their ratio, i.e. $z_t = p_t - d_t$, with $p_t = \log(P_t)$ and $d_t = \log(D_t)$. Specifically, forward recursive or rolling regressions of (2.3) are carried out, providing a sequence of ADF-statistics $ADF_{z,\tau}$ for all margins $\tau = \tau_0, \tau_0 + 1, \ldots, T$. The first regression in the recursive setting hence includes $\tau_0 = \lfloor r_0 T \rfloor$ observations, where $r_0 \in (0, 1)$ is the minimum share of the total sample size T for which (2.3) is estimated. In the rolling regressions the window size is fixed at τ_0.

As outlined above, a bubble is detected when the price series shows explosive behavior while the fundamental series does not. The bubble emergence date $\hat{\tau}_{p,e}$ can then be obtained as the first date τ for which the ADF statistic for the price series ($ADF_{p,\tau}$) exceeds the critical value $cv_{\alpha_T}^{adf}(\tau)$, whereas the collapse date $\hat{\tau}_{p,f}$ is the first subsequent date for which the ADF statistic falls below the critical value

$$\hat{\tau}_{p,e} = \inf_{\tau \geq \tau_0} \left\{ \tau : ADF_{p,\tau} > cv_{\alpha_T}^{adf}(\tau) \right\}, \quad \hat{\tau}_{p,f} = \inf_{\tau \geq \hat{\tau}_{p,e}} \left\{ \tau : ADF_{p,\tau} < cv_{\alpha_T}^{adf}(\tau) \right\}. \tag{2.4}$$

If the dividend series is explosive as well (i.e. $ADF_{d,\tau} > cv_{\alpha_T}^{adf}(\tau)$, for all τ with $\hat{\tau}_{p,e} \leq \tau \leq \hat{\tau}_{p,f}$), no bubble is indicated. If the dividend series turns explosive at $\hat{\tau}_{d,e}$ with $\hat{\tau}_{p,e} < \hat{\tau}_{d,e} \leq \hat{\tau}_{p,f}$, the bubble collapse date $\hat{\tau}_{p,f}$ is reset to $\hat{\tau}_{p,f} = \hat{\tau}_{d,e}$. This algorithm yields a binary indicator series $\mathcal{B}_\tau, \tau = \tau_0, \tau_0 + 1, \ldots T$ with

$$\mathcal{B}_\tau = \begin{cases} 1 & \text{if } ADF_{p,\tau} > cv_{\alpha_T}^{adf}(\tau) \text{ and } ADF_{d,\tau} \leq cv_{\alpha_T}^{adf}(\tau) \\ 0 & \text{else.} \end{cases}$$

Importantly both series can be assessed individually (i.e. $z_t \in \{p_t, d_t\}$), or by their ratio (i.e. $z_t = p_t - d_t$). The literature considers both options but does not

discuss the implications that follow this choice.[7] Rewriting (2.2) in the alternative representation

$$P_t - \frac{1}{r}D_t = \left(\frac{1+r}{r}\right)\sum_{i=1}^{\infty}\left(\frac{1}{1+r}\right)^i E_t\left[\Delta D_{t+i}\right] + B_t, \tag{2.5}$$

however, shows that the difference $P_t - \frac{1}{r}D_t$ is stationary if dividends follow a random walk (with drift), i.e. $D_t \sim I(1)$, and there is no bubble, $B_t = 0$ (Campbell and Shiller, 1987, 1988; Cochrane, 1992). Testing the individual series hence features a larger power against the null hypothesis than the test on the ratio. On the one hand, this implies that testing the series individually can provide a more timely detection of a bubble. On the other hand, this can also imply a more frequent issuance of false positive signals. Similarly, this also applies to the dividend series, which, if found to be explosive, can result in the indicator to signal a bubble collapse. If a false positive signal on the dividend series is issued during the run of the bubble, the indicator would hence signal a bubble's collapse preemptively. Testing the log price-to-dividend ratio can, in contrast, provide complementary information on the relative growth rate of the two series. If both grow exponentially, but prices grow at a faster rate than dividends (indicating increasing imbalances), a unit root test on the ratio can indicate a bubble period that is missed by a test on the individual series due to explosive growth in both series. Therefore, I evaluate both options.

2.2.2 Existing real-time indicators

2.2.2.1 Forward recursive sup ADF Test

First, I apply the original indicator of Phillips et al. (2011) by estimating (2.3) either recursively or by rolling windows. As the authors, I set the share of observations in the initial sample to $r_0 = 0.1$ for the recursive estimation and to $r_0 = 0.2$ for the rolling window estimation. I denote the indicator based on the recursive estimation by $PWY11^{\bullet}_{rec}$, and the indicator based on the longer rolling window estimation by $PWY11^{\bullet}_{rol}$, where $\bullet \in \{i, r\}$ denotes the indicator based on testing the individual

[7] While Phillips et al. (2011) and Harvey et al. (2015) test the series individually, Homm and Breitung (2012), Pavlidis et al. (2014) and Phillips et al. (2015) test the log price-to-dividend ratio.

series ($\bullet = i$), or based on testing their ratio ($\bullet = r$). Therefore, I assess a total of four specifications of the Phillips et al. (2011) indicator.

In order to account for over-rejection in this multiple testing setting, the significance level α_T needs to approach zero asymptotically for an overall significance level of 5% to hold. Correspondingly, the critical values $cv_{\alpha_T}^{adf}(\tau)$ must diverge to infinity. Thus, Phillips et al. (2011) suggest using $cv_{\alpha_T}^{adf}(\tau) = \log(\log(\tau))/100$, yielding significance levels of around 4%. For the rolling $PWY11_{rol}^{\bullet}$ indicator, the critical value is a constant $cv_{\alpha_T}^{adf}(\tau) = \log(\log(\tau_0))/100$. The lag order J is determined by the Akaike Information Criterion (AIC) with $J^{max} = 12$.

2.2.2.2 Generalized sup ADF Test

Phillips et al. (2015) extend the work of Phillips et al. (2011) by not only allowing the end point (here τ_2) to move forward for each recursive regression from $\tau_2 = \tau_0, \ldots, T$, but by also allowing the start point τ_1 for a given τ_2 to vary between all values from 0 to $\max(\tau_2 - \tau_0, 1)$. Thus, the test augments the forward recursive regressions by testing all possible backward extending windows from the current margin τ_2. For a given end point $\tau_2 \in [\tau_0, T]$ and the varying start point $\tau_1 \in [0, \tau_2 - \tau_0]$, the sequence of ADF test statistics is denoted by $\{ADF_{\tau_1}^{\tau_2}\}$. Taking the supremum of this sequence then provides the backward sup ADF statistic for test margin τ_2 denoted by $BSADF_{\tau_2} = \sup_{\tau_1 \in [0, \tau_2 - \tau_0]} \left\{ADF_{\tau_1}^{\tau_2}\right\}$. As in Phillips et al. (2011), emergence and collapse dates are then defined as the first date for which the BSADF statistic exceeds or falls below the respective right-tailed critical value[8]

$$\begin{aligned} \hat{\tau}_e &= \inf_{\tau 2 \in [\tau_0, T]} \left\{\tau_2 : BSADF_{\tau_2} > cv_{\alpha_T}^{bsadf}(\tau_2)\right\}, \\ \hat{\tau}_f &= \inf_{\tau 2 \in [\hat{\tau}_e, T]} \left\{\tau_2 : BSADF_{\tau_2} < cv_{\alpha_T}^{bsadf}(\tau_2)\right\}. \end{aligned} \tag{2.6}$$

Following Phillips et al. (2015) I set the initial sample size to $r_0 = 0.01 + \frac{1.8}{\sqrt{T}}$ and fix the lag order in (2.3) to $J = 1$. Assessing the indicator again both on the individual series and on their ratio, I evaluate two specifications denoted by $PSY15^{\bullet}$, with $\bullet = \{i, r\}$.

[8] Critical values for the $BSADF_{\tau_2}$ statistic are obtained from simulations using the MATLAB code provided by Shu-Ping Shi on her website (`https://sites.google.com/site/shupingshi/home/research`). This code also allows to obtain the $PWY11$ and $PSY15$ indicators.

2.2.2.3 Excessive deviations from HP-filtered trend

Prior to the seminal contribution of Phillips et al. (2011), the literature commonly identified asset price bubbles by evaluating the deviation of the real price series from its one-sided HP-filtered trend. This builds on the assumption that the fundamental series follows a slow-moving trend. As such, this indicator can potentially capture periods in which large deviations of prices from their past trend occur, regardless of the speed of this accumulation process. To compare the novel indicators based on unit root test to the previous literature, I therefore also reevaluate this indicator and identify a bubble if the real stock price deviates from its trend beyond a threshold κ_{hp}, with $\kappa_{hp} = 10\%$ following Assenmacher-Wesche and Gerlach (2010). I adjust the smoothing parameter for estimating the trend component applied in Assenmacher-Wesche and Gerlach (2010) to the monthly frequency, i.e. $\lambda = 8{,}100{,}000$. In addition to the commonly applied recursive method indicated by HP_{rec}, I also assess a rolling estimation with window size $\omega = 96$ denoted as HP_{rol}. These indicators are applied only to the real price series, no to the dividend series. Following the literature, I do not update past estimates of the trend component as new observations arrive but only estimate the τ trend value at each margin based on all available information.

Dating bubbles using the HP-Filter can be criticized along several dimensions. First, setting the threshold and the smoothing parameter is arbitrary and not based on theory. Second, the method can tend to generate more booms in the later part of the sample as the trend estimates becomes less sensitive to new information. Furthermore, the trend component will capture part of the asset price bubble the longer it runs or the more pronounced the appreciation is, thus underestimating the bubble. Nonetheless, the HP-filter may provide more stable estimates than ADF-type tests as it is not affected by short-lived outliers in dividends.

2.2.3 Detecting stock price bubbles in real-time: An illustration

To illustrate the strengths and weaknesses of these real-time monitoring indicators, I compare their provided bubble signals when applied to U.S. stock market data. For this, I use data for the S&P 500 provided by Shiller (2005), focusing on the period from 1975M1 to 2014M7. The underlying fundamental series for stock prices

are the dividends paid out as per the asset pricing equation (2.2).[9] Both the price and dividend series are deflated by the real-time U.S. Consumer Price Index.[10]

Figure 2.2.1 shows the bubble periods in the S&P 500 index as detected by the real-time indicators. The first finding is that all indicators provide strongly varying signals about the emergence and collapse dates of stock price bubbles. This is particularly apparent when comparing the HP indicators to the ADF-type tests of Phillips et al. (2011, 2015). Yet, also some common bubble periods emerge. The pre-"Black Monday" bubble in 1987 is found by three specifications of the Phillips et al. (2011) indicator, one specification of the $PSY15^{i}$ indicator and both HP indicators. Second, the dot-com bubble is detected by all indicators. However, the signals vary widely with emergence being signaled from as early as 1995M5 to as late as 1996M11 and collapse dates ranging from 1999M8 to 2002M4, about seven months prior and more than two years past the peak in real prices. Furthermore, all indicators differ with regard to their stability. While the HP_{rec}, $PWY11^{\bullet}_{rec}$ and $PWY11^{\bullet}_{rol}$ indicators issue stable signals during the course of the dot-com bubble, the $PSY15^{\bullet}$ indicators frequently collapse and re-emerge again.[11] This lack of continuity presents a challenge to policymakers to properly employ countercyclical financial market policies, potentially misjudging the state of the financial cycle entirely. Finally, the $PWY11^{r}_{rol}$ and the $PSY15^{r}$ indicators signal a bubble around the financial crisis period from around 2008M10 to 2009M6. This period is special as the exponential trend in the price-to-dividend ratio comes from a strong decrease of the ratio after an extended period of stability.

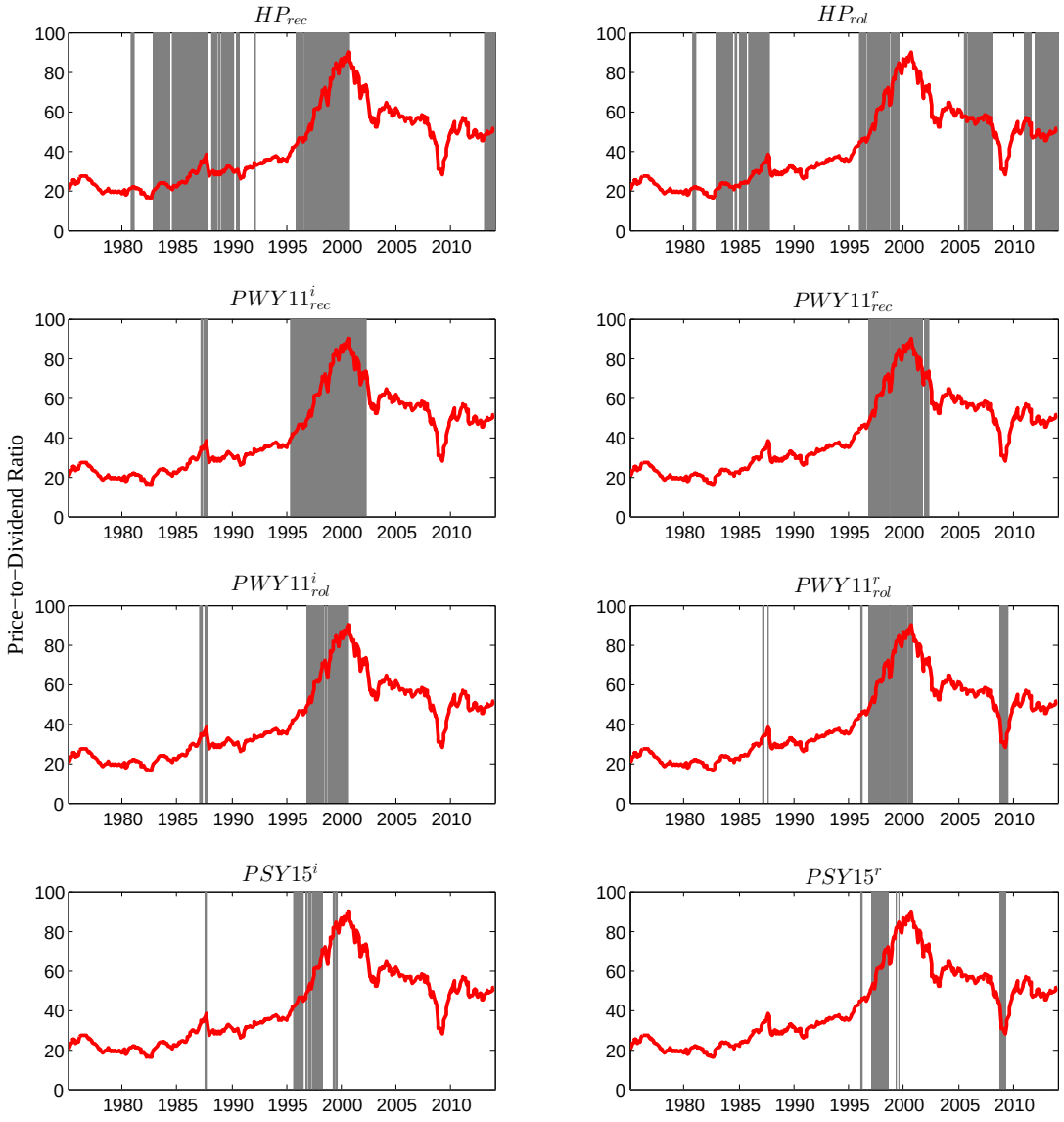
Overall, there are considerable differences between all indicators and their specifications, with few providing continuous bubble periods. In sum, however, the indicators by $PWY11^{i}_{rec}$, $PWY11^{i}_{rol}$ and $PSY15^{i}$ indicators detect the most likely bubble periods, while the HP indicators signal too many bubbles to be plausible.

[9] This data can be obtained from the online supplement of Shiller (2005) available at `http://www.econ.yale.edu/~shiller/data.htm`. I restrict the sample to this period as the forecasting exercise later is carried out for the Great Moderation period starting in the early 1980's.

[10] All real-time data is obtained from the Real-Time Data Set of the Federal Reserve Bank of Philadelphia.

[11] To circumvent this issue, the literature generally suggests bridging bubble periods that are only a few periods apart. While this can be feasible in an ex post analysis, it is not possible in a real-time (forecasting) context.

Figure 2.2.1: Bubble periods in the S&P 500 as detected by individual indicators



Notes: The figure shows the bubble periods in the S&P 500 as detected by the individual indicators. The top panel shows the indicators based on the HP-filter (left: recursive, right: rolling). The bottom three panels show the bubble periods detected by the Phillips et al. (2011, 2015) indicators. The left panels show the indicators when applied to the price and dividend series individually. The right bottom panels show the indicators when applied to the price-to-dividend (PtD) ratio. The solid line plots the PtD ratio.

2.3 A combination approach to real-time bubble detection

The empirical application has emphasized the uncertainty and disagreement among individual indicators in dating the emergence and collapse of asset price bubbles. Yet, the most likely bubble periods are typically detected by all indicators, however, at different times. Hence, I will now outline two strategies to combine the information across indicators, aiming to exploit the individual indicators' different strengths, and by this obtain more robust and reliable signals.

2.3.1 A simple threshold counting approach

First, I aggregate the information content of the individual indicators based on a union of rejections decision rule, following Harvey et al. (2015). The authors combine two ex post tests by Phillips et al. (2011) and Homm and Breitung (2012), obtaining a bubble signal whenever any of the two indicators rejects the null hypothesis of a unit root. To control for the overall size of the power, Harvey et al. (2015) develop critical values for this joint testing strategy.

This approach is however limited when more than two indicators could be used, and when indicators that are not based on statistical testing can provide additional information, e.g. by the HP indicators. Then, inference based on critical values from the joint limiting distribution of all tests is not feasible. Therefore, the first proposed approach shows the merits of indicator combinations by a simple counting approach along different ad-hoc thresholds. Following the spirit of Harvey et al. (2015), however, I identify bubble periods from the union set of all indicators. Thus, the combination indicator with threshold level κ, denoted by $\mathcal{B}^{Comb}_{\tau,\kappa}$, will signal a bubble in period τ if at least κ individual indicators provide a positive signal in τ, i.e.

$$\mathcal{B}^{Comb}_{\tau,\kappa} = \begin{cases} 1 & \text{if } \sum_{\mathcal{B}_\tau \in \mathfrak{B}} \mathcal{B}_\tau \geq \kappa \\ 0 & \text{else} \end{cases}$$

with $\mathfrak{B} = \{PWY11^i_{rec}, PWY11^r_{rec}, PWY11^i_{rol}, PWY11^r_{rol}, PSY15^i, PSY15^r, HP_{rec}, HP_{rol}\}$, and $\kappa = 1, \ldots, K$ with $K = |\mathfrak{B}|$. The smaller κ, the more bubble episodes will be detected, implying an overdetection. The larger κ, the more indicators need to signal a bubble, possibly providing too few signals. A priori, the choice of the optimal threshold κ is not clear. As such, this indicator suffers from the lack of

theoretical guidance about the choice of κ and is sensitive to the included indicators in $\mathfrak{B}$. Nonetheless, this approach provides a straightforward illustration of the potential gains that can be obtained from employing more than one indicator to assess the presence of an asset price bubble in real-time. Furthermore, this approach can be easily implemented at low computational costs, and by this provide a practical solution to aggregate information across indicators. I will explore the best choice for κ for the given set $\mathfrak{B}$ through simulations in section 2.4.

2.3.2 Multiple testing with correlated tests

Addressing the limitations of the counting combinations, I, secondly, propose an indicator based on a statistical combination approach that allows to draw sound inference about the existence of a bubble in real-time and to control the overall size of the indicator. In particular, the proposed monitoring indicator follows the studentized stepwise multiple testing approach of Romano and Wolf (2005). Importantly, this approach also takes any dependence structure of nested, and thereby, correlated tests into account. Yet, while this procedure offers control over the overall size of the combined tests, it comes at the cost of computational intensity. In addition, it requires tests to provide comparable test statistics and is therefore only feasible for a subset of all considered indicators. To outline the usefulness of this indicator and to reduce the computational burden, I will hence combine the three indicators based on the individual series (i.e. $PWY11^{i}_{rec}$, $PWY11^{i}_{rol}$, and $PSY15^{i}$), given the larger power of these indicators compared to testing the price-dividend ratio.[12] The general procedure is, however, applicable to any number of comparable tests. The developed combination indicator will then assure control over the familywise error rate (FWE) of all combined tests, where the FWE is defined as the probability that at least one test falsely rejects the true null hypothesis $H_0 : \delta = 1$. The method will control the FWE to be no greater than $\alpha = 0.05$, at least asymptotically.

This combination approach proceeds as follows. First, I order the $s = 1, \ldots, S$ ADF-type test statistics (here $S = 3$) of all included indicators by their size at each margin τ as $ADF_{r_1,\tau} \geq ADF_{r_2,\tau} \geq \ldots ADF_{r_S,\tau}$, where r_1 denotes the largest, and r_S denotes the smallest statistic.[13] Hence, the most likely significant test statistic

[12] The $PWY11^{\bullet}_{rol}$ demands the largest minimum share of observations to be evaluated. Hence the combination indicator starts at $\tau_0 = \lfloor 0.2T \rfloor$.

[13] Note that for the $PSY15^{i}$ indicator, this is not the ADF but the $BSADF$ test statistic.

$ADF_{r_1,\tau}$ is ordered first. However, it is important to also account for different means and variances of the test statistic. Therefore, I follow Romano and Wolf (2005) and studentize each test statistic where I obtain their means and standard deviations from the bootstrap procedure described below. The joint confidence region for assessing the hypothesis $H_0 : \delta = 1$ with nominal joint coverage probability $1 - \alpha$ then takes the form

$$[\widetilde{ADF}_{r_1,\tau} - c_{1,\tau}, \infty) \times \cdots \times [\widetilde{ADF}_{r_S,\tau} - c_{1,\tau}, \infty) \tag{2.7}$$

where $\widetilde{ADF}_{.,\tau}$ denotes the studentized test statistic, and where the critical value $c_{1,\tau}$ must be obtained from simulations.[14] Romano and Wolf (2005) outline how to use this framework in a stepwise procedure to identify as many tests as possible that reject the null hypothesis. For the application here, it is however sufficient for the largest test statistic $\widetilde{ADF}_{r_1,\tau}$ to reject the null hypothesis and to signal the emergence of a bubble.

To control the α FWE, the critical value c_1 must define the $1 - \alpha$ quantile of the sampling distribution of $\max_{1 \leq s \leq S} \widetilde{ADF}_{r_s,\tau}$. Of course the true probability mechanism P for this distribution is unknown in practice. For this, Romano and Wolf (2005) describe a general algorithm that can be used to obtain an estimate for the required critical value $\hat{c}_1$ that allows to assess the significance of the largest test statistic. I follow this algorithm and first generate a series $z_T^\star$ under the null hypothesis that $z_T^\star$ follows a random walk with drift, with the drift parameter $\mu_z = T^{-1}$. For this series, I then generate $M = 1{,}000$ bootstrap series $z_T^{\star,1}, \ldots, z_T^{\star,M}$.

For each series $z_T^\star$ and all $z_T^{\star,m}$, I compute the test statistics at each margin τ for the three indicators $PWY11^i_{rec}$, $PWY11^i_{rol}$, and $PSY15^i$. After ordering them by size, I denote them by $ADF^\star_{r_1,\tau}, ADF^\star_{r_2,\tau}, ADF^\star_{r_3,\tau}$ for the series $z_T^\star$ and their bootstrap equivalents $ADF^{\star,m}_{r_1,\tau}, ADF^{\star,m}_{r_2,\tau}, ADF^{\star,m}_{r_3,\tau}$. The $M = 1{,}000$ bootstrap draws allow to estimate the mean and standard deviation of the test statistics at each τ. I then use these estimates to studentize the test statistics for the series $z_T^\star$ and $z_T^{\star,m}$. After this standardization step, I retain the largest studentized test statistics of each bootstrap iteration $m = 1, \ldots, M$ and at each margin τ, denoted by $\max^{\star,m}_\tau = \max_{1 \leq s \leq S} \widetilde{ADF}^{\star,m}_{r_s,\tau}$, and obtain the critical value $\hat{c}_{1,\tau}$ as the $1 - \alpha$ empirical quantile

[14] Compare Equation (7) of Romano and Wolf (2005).

of the M values $\max_{\tau}^{\star,1}, \ldots, \max_{\tau}^{\star,M}$.[15] I denote the combination indicator obtained from this application of the Romano and Wolf (2005) algorithm as $\mathcal{B}_{\tau}^{RW}$, signaling a bubble in period τ if $\widetilde{ADF}_{r_1,\tau} > \hat{c}_{1,\tau}$.

Figure 2.B.3 plots the critical values for four different sample sizes obtained from the simulations obtained above (in red). The blue lines provide a fitted curve of the critical values obtained from regressing the critical values for each margin τ on a constant and a term $\log(\tau)$.

To assess the size of this approach, I generate $S = 200$ series $z_T^{\star,s}$ under the null hypothesis and perform the above described algorithm on each series. Table 2.3.1 shows the results for different sample sizes T, together with the finite sample sizes for the individual indicators used in the combination. The table reveals an upward deviation from the nominal size of $\alpha = 0.05$ for all indicators and sample sizes with the exception of the $PWY11_{rol}^{\bullet}$ for $T = 200$. The $PSY15^{\bullet}$ indicator appears to suffer most from overdetection in these short samples. In contrast to this, the combination indicator is relatively well aligned with the nominal size, yet overdetection becomes more prominent the longer the sample. This is due to the additional multiple testing that emerges from the iterative application of the algorithm, yet this issue appears to be of second-order relevance.[16] This is also supported by Figure 2.B.4 which shows the share of false positive signals across all simulations over the sample. Overall, this share not trending over time. All in all, I thus conclude that the algorithm provides critical values that control the nominal size of the test.

2.3.3 Empirical illustration

Figure 2.3.2 displays the stock price bubble periods detected by the combination indicators $\mathcal{B}_{\tau,\kappa}^{Comb}$ for $\kappa = 2, \ldots, 6$ and the $\mathcal{B}_{\tau}^{RW}$ indicator (bottom right panel). For $\kappa = 2$ (and $\kappa = 1$ not depicted here), several periods that do not feature a prominent increase in the price-to-dividend ratio are classified as bubbles, similar to the HP indicators. Furthermore the dot-com bubble period extends until 2001M8, far beyond the price crash in 2000M3. For $\kappa = 3, 4, 5$, bubbles are detected only

[15] I summarize this algorithm in the Appendix 2.A. This algorithm is analogous to the first iteration of Algorithm 4.2 of Romano and Wolf (2005).

[16] I further note that computing the critical value as the $1 - \alpha$ quantile of the largest studentized test statistics along all margins τ, i.e. $\max_{\tau_0}^{\star,1}, \ldots, \max_{T}^{\star,1}, \ldots, \max_{\tau_0}^{\star,M}, \max_{T}^{\star,1}$, does not reduce overdetection.

Table 2.3.1: Size of the combination indicator $\mathcal{B}_{\tau}^{RW}$

	T			
Indicator	**200**	**300**	**480**	**624**
$PWY11^{\bullet}_{rec}$	0.0605	0.0605	0.0544	0.0533
$PWY11^{\bullet}_{rol}$	0.0481	0.0505	0.0522	0.0537
$PSY15^{\bullet}$	0.0601	0.0684	0.0653	0.0641
$\mathcal{B}_{\tau}^{RW}$	0.0525	0.0560	0.0564	0.0567

Notes: The table shows the size of the three individual tests and of their combination following the Romano and Wolf (2005) algorithm for four different sample sizes and a nominal size of $\alpha = 0.05$.

around 1987M3 before the 1987M10 crash and around 1995M12 for the dot-com bubble. Also, the collapses are detected on time. Increasing κ further deteriorates the emergence detection, and the dot-com bubble collapse is dated well before the ultimate peak in prices.

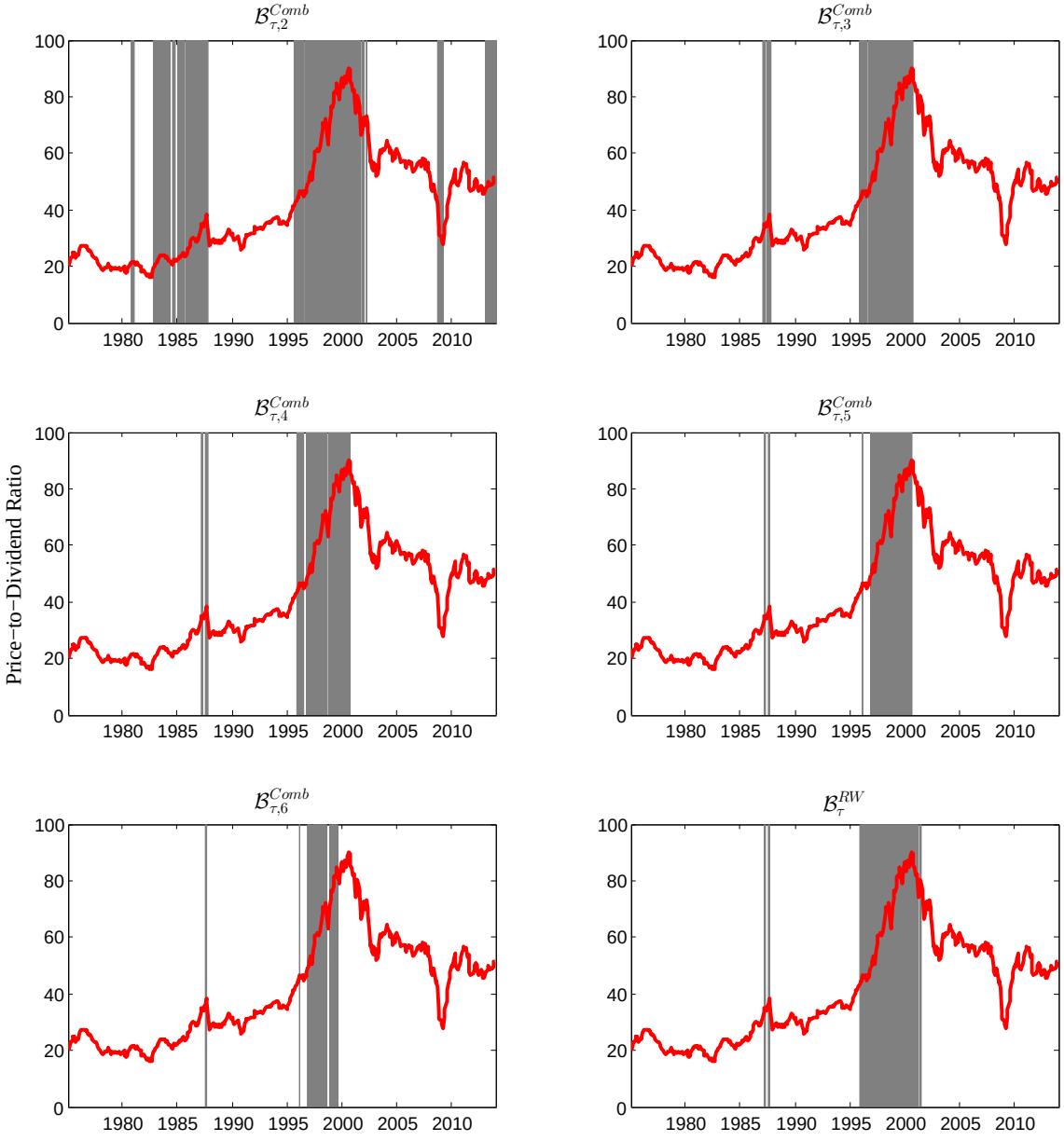
The econometric combination $\mathcal{B}_{\tau}^{RW}$ combining the $PWY11^{i}_{rec}$, $PWY11^{i}_{rol}$ and $PSY15^{i}$ indicators mirrors the findings of the most accurate counting combinations. It signals the period before the 1987 "Black Monday" crash and detects the dot-com bubble early in 1995M11. In contrast to the combinations with $\kappa \geq 3$, the collapse, however, is detected only in 2001M3, one year after the peak in 2000M3. Yet, in contrast to the counting combinations with $\kappa > 3$, the econometric combination performs better in terms of signal stability with no collapse and re-emergence during the run of the dot-com bubble.

All in all, the combination indicators with $3 \leq \kappa \leq 5$, and the $\mathcal{B}_{\tau}^{RW}$ indicator appear to provide reasonable and balanced bubble signals, capturing the most prominent stock price bubbles early after emergence, and avoiding the likely false positive signals of some of the individual indicators. Moreover, these indicators provide more stable signals during the run of the dot-com bubble than the $PSY15^{\bullet}$ indicators.

2.4 Finite sample power and accuracy

I now evaluate the finite sample power properties of this extensive set of indicators in a controlled experiment. For this, I generally follow the simulation set-up of Phillips

Figure 2.3.2: Bubbles periods in the S&P 500 as detected by combination indicators



Notes: The figure shows the bubbles periods in the S&P 500 as detected by the ad-hoc counting combination indicators $\mathcal{B}^{Comb}_{\tau,\kappa}$ for different threshold levels κ, and by the econometric combination $\mathcal{B}^{RW}_{\tau}$ (bottom right panel). The solid line plots the PtD ratio.

et al. (2015) but I not only simulate a price but also a dividend series in order to study the effect of testing either both series individually or their log-ratio.

2.4.1 Bubbles as mildly explosive processes

I study the finite sample power properties of all indicators when the price series follows a mildly explosive process that is capable of generating a fixed number of bubbles over a specified sample length. For the single bubble case, this process follows

$$z_t = z_{t-1} I\{t < \tau_e\} + \delta_T z_{t-1} I\{\tau_e \leq t \leq \tau_f\} + \left(\sum_{k=\tau_f+1}^{t} \varepsilon_k + z^*_{\tau_f} \right) I\{t > \tau_f\} + \varepsilon_t I\{t \leq \tau_f\}, \tag{2.8}$$

where $\delta_T = 1 + cT^{-\alpha}$ with $c > 0$ and $\alpha \in (0, 1)$, $\varepsilon_t \sim (0, \sigma^2)$ and $z^*_{\tau_f} = z_{\tau_e} + z^*$ with $z^* = O_p(1)$. Until bubble emergence and after collapse, the process is thus characterized by a random walk. During the bubble period, the process is explosive with an expansion rate $\delta_T > 1$. After the collapse, the process returns to the pre-bubble value plus a small perturbation z^*.

As Phillips et al. (2015) emphasize, it is crucial for bubble tests to be able to "restart" after an initial bubble was detected and collapsed. Therefore, I run a second simulation that features two explosive and collapsing processes. This two-bubble scenario is described accordingly with two mildly explosive bubble periods (characterized by the same growth rate d_T) and random walk processes before, in between, and after the respective bubble periods

$$\begin{aligned} z_t =& z_{t-1} I\{t \in N_0\} + \delta_T z_{t-1} I\{t \in B1 \cup B2\} + \left(\sum_{k=\tau_1 f+1}^{t} \varepsilon_k + z^*_{\tau_1 f} \right) I\{t \in N_1\} \\ &+ \left(\sum_{l=\tau_2 f+1}^{t} \varepsilon_l + z^*_{\tau_2 f} \right) I\{t \in N_2\} + \varepsilon_t I\{t \in N_0 \cup B_1 \cup B_2\}, \end{aligned} \tag{2.9}$$

where $N_0 = [1, \tau_{1e})$, $B_1 = [\tau_{1e}, \tau_{1f}]$, $N_1 = (\tau_{1f}, \tau_{2e})$, $B_2 = [\tau_{2e}, \tau_{2f}]$, $N_2 = (\tau_{2f}, T]$. Extending the work of Phillips et al. (2015), I further generate a sequence of dividends that are assumed to follow a random walk with drift, where the drift parameter

$\mu = 0.38$ matches the sample estimate for the S&P 500 dividend series. The parameters in (2.8) and (2.9) take the values $c = 1$ and $\alpha = 0.6$ as in Phillips et al. (2015). The processes are initialized with $z_0 = 100$, and are restricted to remain positive throughout. The variance of the disturbance ε_t is matched to the sample standard deviation of the S&P 500 prices series, i.e. $\sigma = 8.94$. I evaluate the finite sample properties over 5,000 simulations.

2.4.2 Simulation results: single bubble process

I present the results for the indicators' finite sample power properties and accuracy for the single bubble case in Tables 2.C.2 to 2.C.8. First, to explore the effects of bubble location and duration, as well as the specification choices for the Phillips et al. (2011, 2015) indicators on their detection abilities, I assess the *number of identified continuous bubble episodes* in Tables 2.C.2 and 2.C.3. These results suggest that most indicators identify too many bubble periods. For this reason, I then assess the indicators' accuracy along six dimensions. First, I present the *share of false decisions* (Type I and II errors) in Table 2.C.4, evaluating each indicator's overall accuracy. For policymakers type I and II errors pose a natural trade-off. Depending on the policymakers' preferences or the policy under consideration, *missing a bubble* may be more costly than receiving a *false positive signal*, or vice versa. Therefore, I evaluate both sources of error separately in Tables 2.C.5 and 2.C.6. Furthermore, as highlighted in the empirical application, another source of concern are frequent collapse and re-emergence signals during a bubble's run. I assess the indicators' *stability* in Table 2.C.7, displaying the share of missing signals in-between the earliest emergence and the latest collapse dates provided by each indicator during the true bubble episode. Finally, I explore the *delay until emergence and collapse* is detected by each indicator in Table 2.C.8.

Overall, my findings emphasize that higher success rates of detecting the true bubble or a higher signal stability come at the cost of more false positive signals, and a longer delay until a bubble collapse is detected. When weighing these costs and benefits, the $PWY11^i_{rec}$ and the $PSY15^i$ indicators emerge as the most promising real-time monitoring indicators from the set of individual indicators. Nonetheless, I find that the most promising combining indicators provide considerable gains against these best individual indicators with regard to the accuracy of the signaled emer-

gence and collapse dates, and for the stability of the signal during the course of the bubble. Here, the results are in line with the empirical application, supporting both the econometric combination and the simple counting combinations with a threshold choice of $\kappa = 3$ or $\kappa = 4$.

2.4.2.1 Average number of detected bubble episodes

I begin by assessing the average number of continuous bubble episodes signaled by each indicator in Table 2.C.2 for different sample lengths T, emergence dates τ_e and bubble durations τ_d. Here, I find that the HP_{rec}, HP_{rol} and $PWY11^{\bullet}_{rol}$ indicators deviate most strongly from capturing the true one-bubble process by providing the highest number of bubble episodes (between 2.2-3.9 bubbles on average), suggesting either frequent on-off signals during the bubble's run, or false positive signals.[17] In contrast to this, the average detection rates for the $PWY11^{\bullet}_{rec}$ and $PSY15^{\bullet}$ indicators are more closely aligned to the true one-bubble process. Yet, all three parameters of the bubble process (sample length, location, and duration) determine which indicator is the most accurate by this criterion.

The effect of different bubble *emergence dates* on the indicators' detection abilities is relatively modest for most indicators with the exception of the $PSY15^{\bullet}$ indicators. Similar to Phillips et al. (2015), I find that these indicators show lower detection rates when the bubble starts early in the sample (all else being equal). More important than the emergence date is the *duration* of the bubble. With the exception of the HP indicators, the number of detected bubbles increases with bubble duration.[18] Adding to the previous literature, I find that the *choice of the tested series* matters strongly for the performance of the Phillips et al. (2011, 2015) indicators. As outlined in section 2.2.1, I find that all these indicators detect fewer bubbles when the (log) price-dividend ratio is assessed compared to when both series are tested individually. All in all, the results suggest that the $PWY11^{\bullet}_{rec}$ and $PSY15^{\bullet}$ mirror a one-bubble

[17] For the HP indicators, overdetection could potentially be addressed by increasing the threshold κ_{hp}. I do not seek to optimize this indicator class but evaluate the indicator as employed by previous studies (Adalid and Detken, 2007; Assenmacher-Wesche and Gerlach, 2010).

[18] In order to rule out short-lived bubbles, Phillips et al. (2011, 2015) impose an additional ex post constraint on the minimum length of the bubble. For the real-time monitoring procedure of interest to this paper, such an ex post adjustment is, however, not feasible. Therefore, my results show higher detection rates than the results of Phillips et al. (2015).

process most closely, yet this criterion does not provide any evidence for the signals' accuracy of detecting the true bubble.

In the bottom panel of Table 2.C.2, I evaluate if the combination indicators can provide bubble signals that are less sensitive to the specific bubble process at play. Generally, I find that the simple combination with $\kappa = 5$ provides reasonable signals of around one bubble on average. Furthermore, the performance of the counting combinations is, in contrast to the $PSY15^{\bullet}$ indicator, not affected by the sample size or the bubble location. This remains an issue for the $\mathcal{B}_{\tau}^{RW}$ indicator. Bubble duration on the other hand still plays a key role in how many bubbles are detected for both combination approaches.

2.4.2.2 Frequency distribution of bubble signals

The gains from the combination indicators in reducing the sensitivity to the location of the bubble become more visible in Table 2.C.3, showing the frequency distribution of bubble signals for two simulations with $T = 200$, $\tau_d = 20$ and two different emergence dates τ_e.[19] Generally, the results are in line with the results discussed above with the $PWY11_{rol}^{\bullet}$ and the HP indicators signaling too many bubbles, and the $PWY11_{rec}^{\bullet}$ and the $PSY15^{\bullet}$ indicators detecting the true number of one bubble more frequently. However, the impact of bubble location becomes apparent. Here, the $PWY11_{rec}^{i}$ indicator most frequently detects the true number of one bubble in 47.2% of simulations when the bubble occurs early in the sample. When the bubble starts later, however, the $PSY15^{i}$ indicator performs best with a detection accuracy of 52%.

As documented in the lower panel the counting combination indicators are less sensitive to the location of the bubble with the $\mathcal{B}_{\tau,4}^{Comb}$ indicator detecting the correct number of bubbles in over 50% of the cases regardless of the emergence date. Hence, these indicators can provide a useful insurance against bubble location. The $\mathcal{B}_{\tau}^{RW}$ combination, in contrast, performs better the later the bubble emerges, detecting the true number of one bubble in 46.8% (for $\tau_e = 40$) and 59.1% (for $\tau_e = 120$) of all simulations.

[19] Results for detection rates when altering τ_d and T are in line with the above discussion.

2.4.2.3 Type I and type II error rates

To assess the accuracy of all indicators in detecting the true bubble, I first assess the overall share of false decisions in Table 2.C.4. This table thus shows how prone each indicator is to Type I errors (missing a bubble) and Type II errors (issuing a false signal). Among the individual indicators, the HP indicators are the least reliable with error rates of around 15% to 20%. Similarly, the $PWY11^{\bullet}_{rol}$ indicators feature high error rates for longer bubble periods. In contrast, the $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators usually achieve considerably lower error rates of around 10%. Their counterparts $PWY11^{r}_{rec}$ and $PSY15^{r}$, however, show higher error rates of up to 17.7%. Again, the indicators' detection abilities depend on the characteristics of the bubble as discussed above.

In contrast to the individual indicators, the combination indicators' error rates are less sensitive to the bubble process. Here, I find that simple combinations with $\kappa = 3$ or $\kappa = 4$ and the econometric combination $\mathcal{B}^{RW}_{\tau}$ perform best, and generally outperform the best-performing individual indicator. Hence, these results emphasize the gains that both combination approaches can provide. In the following I will condition the discussion on these overall error rates. In particular, I will discard the results for combination indicators with $\kappa < 3$ and $\kappa > 6$, and I will further focus on the $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators in the discussion of the best-performing individual indicators.

To assess the source of error in more detail, I display the share of false signals (Type II error) to the total number of signals in Table 2.C.5. These results emphasize that both the HP as well as the $PWY^{\bullet}_{rol}$ indicators suffer most from false positive signals, with Type II error rates as high as 76.9%. This is particularly problematic for simulations with short bubbles. In contrast, the more parsimonious $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators considerably reduce the probability of falsely signaling a bubble. Yet, further gains can again be provided by the combination indicators. Here, the combinations with $\kappa = 4$ and $\kappa = 5$ always outperform all individual indicators. For longer bubbles, or bubbles that occur later in the sample, the econometric combination $\mathcal{B}^{RW}_{\tau}$ achieves similarly low error rates as these best counting combinations. Furthermore the $\mathcal{B}^{RW}_{\tau}$ indicator outperforms all individual indicators for almost all bubble specifications. Hence, these combinations offer a clear improvement over the individual indicators in providing accurate positive sig-

nals and are considerably less sensitive to the location and duration of the bubble process.

Importantly, this does not come at the cost of failing to detect true bubble periods as documented in Table 2.C.6. Here, I display the period-by-period probability of a bubble occurrence given that no signal is issued. Obviously, this share is the lowest for indicators that issue the most signals, in particular the HP indicators.[20] Among the indicators with the best overall error rates, the $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators perform best. Again, later bubbles are more accurately detected by the $PSY15^{i}$ indicator. Furthermore, the combination indicators provide possible gains to the individual indicators. Here, a policymaker with a stronger preference of not missing a bubble could choose a combination with $\kappa = 3$. However, also the $\mathcal{B}^{Comb}_{\tau,4}$ and $\mathcal{B}^{RW}_{\tau}$ indicators frequently outperform the best individual indicators, and provide the best balance between Type I and Type II error rates independent on the unknown characteristics of the bubble process.

2.4.2.4 Indicator instability: frequency of on-off signals

While false positive signals and a failure to detect a true bubble are a major concern for the feasibility of any indicator, the empirical application has also raised questions about indicator instability. This source of error is a form of Type I error, with an indicator failing to signal a true bubble despite a previous correct positive signal. I evaluate this issue in Table 2.C.7. Here, I show the share of missed bubble periods between the earliest and latest positive signal during run of the true bubble.[21] The HP and the $PWY11^{i}_{rec}$ indicators provide the most stable signals while the $PWY11^{\bullet}_{rol}$ indicators suffer most from frequent on-off signals. As can be expected, instability is of stronger concern the longer a bubble runs.

Among the counting combination indicators, a lower threshold κ leads to more stable bubble signals. However, the best indicators $\mathcal{B}^{Comb}_{\tau,3}$ and $\mathcal{B}^{Comb}_{\tau,4}$ do not seem to enhance indicator stability, especially for longer bubbles. In contrast, the econometric combination $\mathcal{B}^{RW}_{\tau}$ provides stable signals and outperforms the $PWY11^{i}_{rec}$ indicator for early bubbles in a short sample.

[20] However, it becomes apparent that the high share of Type II errors of the $PWY11^{\bullet}_{rol}$ indicators does not translate into higher detection rates of the true bubble.

[21] If a bubble is missed completely, this simulation is ignored.

2.4.2.5 Signal delay

Finally, I evaluate the average delay for each individual indicator until emergence and collapse of the true bubble are signaled in Table 2.C.8. From the left panel, it becomes apparent that the delay until emergence is signaled can be substantial. When applied to monthly data, all indicators frequently miss the bubble over the course of the first year. The earliest signals are generally given by the indicators that signal the most bubbles, most notably the HP indicators. Again, for the overall most accurate $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators, bubble location is a decisive factor. For earlier bubbles, the $PWY11^{i}_{rec}$ features higher detection rates, while the $PSY15^{i}$ performs better for later bubbles. While the combination with $\kappa = 3$ provides the earliest bubble signals, the more accurate combination indicator with $\kappa = 4$ and the econometric combination $\mathcal{B}^{RW}_{\tau}$ are also competitive compared to the $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators in detecting bubble emergence on time.

In contrast to the delay until emergence, the right panel shows that most indicators detect the collapse of a bubble almost immediately or preemptively. Here, it is always the $PSY15^{i}$ indicator that provides the most immediate collapse warnings. Also, the right lower panel shows that combination indicators with $3 \leq \kappa \leq 5$ and the $\mathcal{B}^{RW}_{\tau}$ indicator are competitive compared to the $PSY15^{i}$ indicator.

All in all, the results show that the $PWY11^{i}_{rec}$ and $PSY15^{i}$ indicators detect asset price bubbles reasonably well, balancing the number of false positive signals and missed bubbles. Also both indicators detect the collapse date on time. The issue of false alarms and collapse date inaccuracy, in contrast, are main concerns with the $PWY11^{\bullet}_{rol}$ and the HP indicators. Nonetheless, the HP indicators provide valuable additional information as it suffers least from instability and detects the emergence of a bubble with the shortest delay. As such, combination indicators that make use of HP indicators can employ the complementary strengths of these available indicators. Yet, also the econometric combination approach $\mathcal{B}^{RW}_{\tau}$ frequently provides considerable gains to simply assessing single indicators. Overall, I find counting combination indicators that require $\kappa = 3$ or $\kappa = 4$ individual indicators to signal a bubble simultaneously and the $\mathcal{B}^{RW}_{\tau}$ indicator to provide more reliable bubble signals and, importantly, to be less sensitive to the characteristics of the bubble process than the best performing individual indicators.

2.4.3 Simulation results: two collapsing bubbles

Phillips et al. (2015) show that the $PWY^{\bullet}_{rec}$ indicator has troubles to restart after a first bubble has collapsed and frequently fails to detect a second bubble in the sample. This is strongly improved by the $PSY15^{\bullet}$ indicator. I confirm this for a simulation with $T = 200$ and two bubbles of equal length $\tau_d = 20$ that emerge at $\tau_{1e} = 40$ and $\tau_{2e} = 120$ in Table 2.C.9. The table presents the average number of separate bubble periods detected by the indicators, the respective frequency distribution, as well as the overall error rate (column labeled I+II), the share of false signals (Type II error) and the probability of a bubble occurring if no signal is issued (Type I error).

As in the single bubble case, the HP and the $PWY11^{\bullet}_{rol}$ indicators detect too many bubbles on average and signal more than two bubbles in over half of the simulations. The $PSY15^{\bullet}$ and the $PWY11^{\bullet}_{rec}$ indicators, on the other hand, frequently detect fewer than two bubbles. However, the $PSY15^{i}$ indicator dominates all other individual indicators. This translates further to the error rates. The best signals for the multiple bubble case are, however, again provided by the combination indicators. Here, the combination indicators $\mathcal{B}^{Comb}_{\tau,3}$, $\mathcal{B}^{Comb}_{\tau,4}$, and the econometric combination $\mathcal{B}^{RW}_{\tau}$ generally detect around two bubbles and also minimize the overall and the Type II error rate. Hence, this case emphasizes that indicator combinations provide considerable gains in signal accuracy and reliability compared to the signals provided by individual indicators.

2.5 Forecasting output with bubble indicators

In order to demonstrate the practical relevance of obtaining accurate bubble signals for stock price bubbles, I explore the predictive content of such indicators for forecasting real activity. These bubble indicators are a natural candidate for forecasting real economic developments for two reasons. As demonstrated by the global financial crisis, an asset price bubble can reduce real economic activity following its burst due to financial instability, and the deleveraging of the private and public sector. Yet, asset price bubbles may also directly affect output during the boom phase. Rising asset prices alleviate credit constraints for firms and households, thereby stimulating investment and consumption. Bubble periods can thus intensify regular business

cycle movements both through the boom and the bust (see e.g. Kiyotaki and Moore, 1997 and Bernanke et al., 1999).[22]

2.5.1 Forecast specification

I generally follow Assenmacher-Wesche and Gerlach (2010) who evaluate the predictive content of the HP_{rec} indicator for forecasting output growth. Yet, I extend their paper not only by enlarging the set of bubble indicators, but also by taking the real-time dimension of all variables into account. In contrast to common predictor variables used to forecast output, stock prices and hence the stock price bubble indicators are available in real-time. As Assenmacher-Wesche and Gerlach (2010), I demonstrate the potential of the bubble indicators by predicting industrial production (IP) by single-equation models including variables that are documented as frequently showing significant predictive abilities (Stock and Watson, 2003). In particular, the benchmark model against which I evaluate the marginal predictive content of all bubble indicators includes inflation, the unemployment rate, the effective funds rate and the term spread between 10-year and 3-months government bond yields. The forecast model hence takes the specification

$$y_{\tau+h}(t) = x'_{m,\tau}(t)\beta_m + \varepsilon_{m,\tau+h}(t), \quad \varepsilon_\tau \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \tag{2.10}$$

where $y_\tau(t)$ denotes the vintage t observation of y at time τ, with $t \geq \tau$.[23] Hence, for variable y published with a lag of q months, the observations available at time t are $\{y_\tau(t)\}_{\tau=1}^{t-q}$. The full sample then includes observations for the target variable

[22] The negative real effects are likely to be largest for credit financed asset bubbles, in particular on housing markets (Borio and Lowe, 2004; Schularick and Taylor, 2012). A natural extension is thus to investigate the predictive content of credit-financed housing bubbles. I extend this forecast exercise in Chapter 3 where I enlarge the space of competitor models, and also explore the predictive content of these indicators applied to the U.S. housing market.

[23] I evaluate forecasts both with a fixed-lag length of all predictor variables and a specific-to-general lag length selection following Herwartz (2010). As the model with no lags provided the best benchmark forecast, I only show these results. The forecast evaluation is carried out against the final revised values $y_{\tau+h}(T)$ and the evaluation criterion for forecast accuracy is the root mean square prediction error (RMSPE). Direct h-step forecasts for y are obtained for each margin $t = R, \ldots, T-h-q$, since an iterative forecasting procedure would require predictions up to $h-1$ for all variables in x. While this may be feasible for all variables included in the benchmark model, forecasting the path of the bubble is not.

y and the vector of predictors $\{x'_\tau(t)^t_{\tau=1}\}^T_{t=R}$, where x_k-specific publication lags and number of revisions are taken into account.

The benchmark model only includes the classical predictors discussed above. This benchmark model is then augmented by the bubble indicators individually. When an indicator does not signal a single bubble episode in the estimation window $\tau = 0, \ldots, t-h-q$ for forecasting margin t, the benchmark forecast is applied at that margin. That is, the benchmark and the augmented model are identical until the first time that at least one bubble period is signaled. The first forecast is carried out in 1997M12.

2.5.2 Forecast results

I evaluate the predictive accuracy of all models in Table 2.C.10. The benchmark model's RMSPE is given in the first row, followed by the relative gains and losses of the augmented models subsequently. A value below one indicates a better predictive accuracy of the augmented model compared to the benchmark. The best forecast model for each horizon and in each indicator class is highlighted in bold, and the significance of each model's gain or loss is assessed by the Clark and McCracken (2009) test for equal predictive ability for real-time data.

Overall, several indicators for stock price bubbles provide considerable gains compared to the benchmark model, yet these gains differ across the indicators and forecast horizons. For short-run forecasts up to $h = 6$ months, most indicators except for the $PWY11^r_{rol}$ and the $\mathcal{B}^{Comb}_{\tau,1}, \mathcal{B}^{Comb}_{\tau,2}$ outperform the benchmark. Here, the combination indicators with $3 \leq \kappa \leq 5$ and the $PWY11^i_{rol}$ indicator perform best. These best-performing indicators accurately detect the collapse of the dot-com bubble in contrast to their competitors. However, also the individual indicators $PWY11^i_{rec}, PSY15^i, HP_{rec}$, and the econometric combination $\mathcal{B}^{RW}_\tau$ provide considerable and significant gains to the benchmark model. All indicators that outperform the benchmark detect the bubble up to the 1987 stock market crash and the dot-com bubble, but generally do not provide a false positive signal during the global financial crisis. In contrast, indicators that either miss the 1987 bubble ($PWY11^r_{rec}$) or provide a signal around 2009 ($PWY11^r_{rol}, PSY15^r$) frequently perform worse than the benchmark for these horizons. These results mostly transfer to longer horizons with $h > 6$ with the exception of the deterioration in the performance of the HP_{rec} indica-

tor which provides the most bubble signals throughout the sample. In contrast, the parsimonious $PSY15^i$ indicator performs better for longer horizons and provides the most accurate forecasts overall. Likewise the combinations $\mathcal{B}_{\tau,3}^{Comb}, \ldots, \mathcal{B}_{\tau,6}^{Comb}$ continue to outperform the benchmark.

All in all, these results emphasize the importance of accurate signals about the emergence and collapse of asset price bubbles on stock markets, highlighting their prospective value for policymakers. False signals during normal times lower the predictive value of such indicators. Likewise, indicators that fail to capture the collapse of a bubble on time do not help to predict real economic developments following the crash. I confirm these findings in an extended forecast experiment in Chapter 3, where I show that these predictive gains hold even when comparing the indicators' predictive ability against a large set of macroeconomic and financial variables commonly used to forecast real economic activity.

2.6 Conclusion

Recently, Phillips et al. (2011, 2015) have developed promising new indicators that allow to monitor asset markets for the emergence and collapse of bubbles in real time. In this paper, I review these existing approaches and propose two methods to combine the signals from these individual real-time indicators. Applying the existing individual indicators to U.S. stock market data, I highlight that these indicators provide strongly heterogeneous bubble signals. In particular, these indicators frequently suffer from missing true bubble episodes or providing false positive signals, and from signaling collapse and reemergence during the run of a bubble. Yet, when a comprehensive picture is taken, most indicators signal the run-up to the 1987 stock market crash and the dot-com bubble in the late 1990's.

Building on these findings, I develop two combination indicators to aggregate the individual bubble signals in real time. In particular, the first class of combination indicators counts the number of simultaneous signals received from the individual indicators, and requires this number to exceed a specified threshold. While being easy to implement, this approach can, however, not control the overall size of the combination indicator and is sensitive to the threshold choice. Therefore, I develop a second indicator that combines different specifications of the Phillips et al. (2011,

2015) right-tailed unit root tests by means of a stepwise multiple testing procedure that accounts for the correlation of all individual tests.

Through simulations I show that the best choice of an individual indicator is sensitive to the number of bubbles in the sample, their locations, as well as their duration. In contrast, the proposed combination indicators are considerably less sensitive to the characteristics of the bubble process. Importantly, these gains do not come at the cost of lower accuracy. Instead, I show that both the best counting combination indicator as well as the econometric combination outperform the best individual indicators by lowering the overall error rates of the provided signals.

Finally, I demonstrate the value of receiving accurate real-time information about emerging and collapsing asset price bubbles by showing that the most accurate indicators provide considerable gains for predicting output growth for up to two years. Hence, these indicators can enrich the information set of policymakers and of other practitioners.

Appendix

2.A Romano and Wolf (2005) algorithms for combination indicator

Algorithm 1: Combination indicator based on multiple testing with correlated tests: At each margin τ

1. Order the studentized ADF-type test statistics of $PWY11^r_{rec}$, $PWY11^r_{rol}$, and $PSY15^r$ by absolute size, labeling them r_1 for the largest, and r_S for the smallest.

2. Reject the null hypothesis $H_0 : \delta = 1$ if $\widetilde{ADF}_{r_1,\tau} > \hat{c}_{1,\tau}$.

3. If H_0 is rejected, set the indicator $\mathcal{B}^{RW}_\tau = 1$, else $\mathcal{B}^{RW}_\tau = 0$.

Algorithm 2: Computation of the critical value $\hat{c}_1$ via the bootstrap:
Generate a series $z^\star_T$ under the null hypothesis that $z^\star_T$ follows a random walk with drift.

1. Generate $M = 1,000$ bootstrap series $z^{\star,1}_T, \ldots, z^{\star,M}_T$.

2. For each $z^{\star,m}_T$, compute the test statistics associated with the indicators, denoted by $ADF^{\star,m}_{r_1,\tau}, ADF^{\star,m}_{r_2,\tau}, ADF^{\star,m}_{r_3,\tau}$.

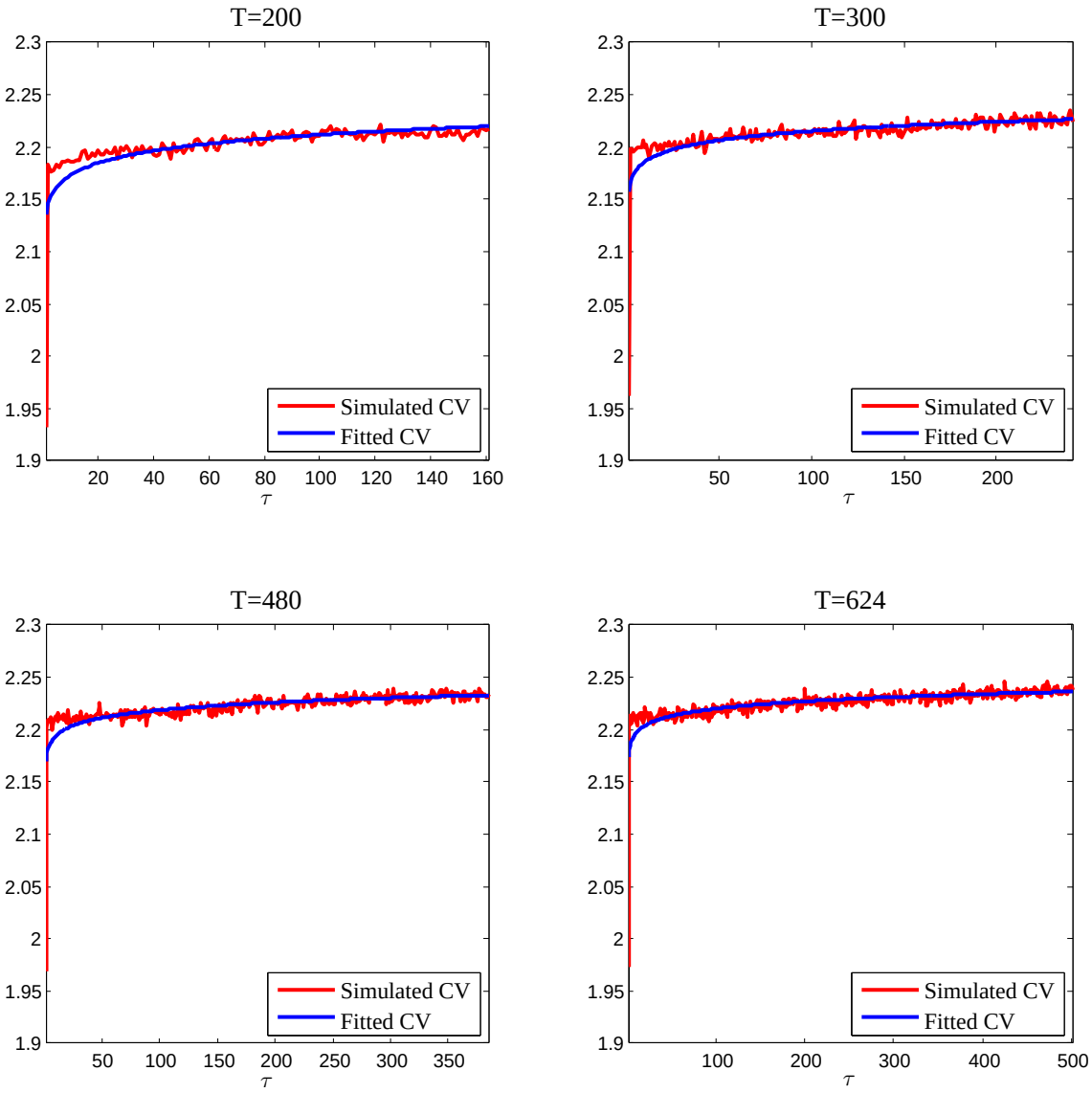
3. Obtain estimates for the mean and standard deviation of each test statistic and studentize test statistics.

4. Obtain largest studentized test statistic for each τ and each bootstrap iteration m as $\max^{\star,m}_\tau = \max_{1\leq s\leq S} \widetilde{ADF}^{\star,m}_{r_s,\tau}$.

5. Compute $\hat{c}_{1,\tau}$ as the $1-\alpha$ quantile of $\max^{\star,1}_\tau, \ldots, \max^{\star,M}_\tau$.

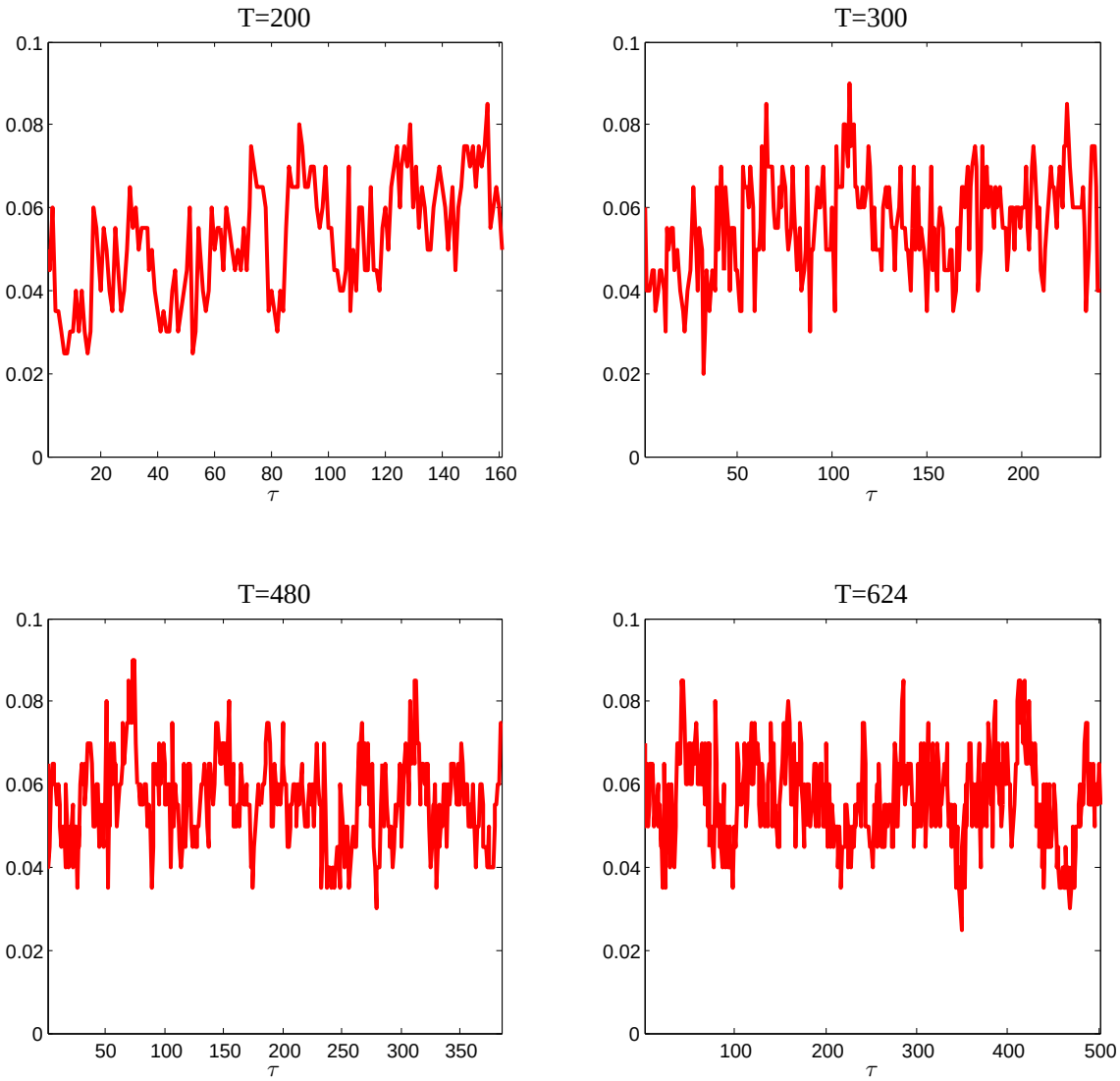
2.B Figures

Figure 2.B.3: Critical values for the $\mathcal{B}_{\tau}^{RW}$ combination indicator



Notes: The figure shows the critical values for the $\mathcal{B}_{\tau}^{RW}$ combination indicator for different sample sizes T. The red line shows the critical values obtained from the bootstrap simulation described in section 2.3.2. The blue line provides a fit line from regressing the simulated critical values on a constant and $\log(\tau)$.

Figure 2.B.4: Period-by-period size of the $\mathcal{B}_\tau^{RW}$ combination indicator



Notes: The figure shows the size (the frequency of a false positive signal) of the $\mathcal{B}_\tau^{RW}$ indicator across time τ for four different sample sizes T.

2.C Tables

Table 2.C.2: Average number of detected bubbles

Indicator	**Average number of detected bubbles**					
T	200	200	200	200	300	300
τ_e	40	40	120	120	60	180
τ_d	20	40	20	40	30	30
HP_{rec}	2.49	2.19	2.43	2.23	3.33	3.08
HP_{rol}	2.80	2.77	2.53	2.25	3.83	3.91
$PWY11^{i}_{rec}$	1.20	1.56	1.26	1.65	1.16	1.19
$PWY11^{r}_{rec}$	1.08	1.53	1.17	1.53	**0.93**	**1.01**
$PWY11^{i}_{rol}$	3.09	3.45	3.06	3.51	2.87	2.86
$PWY11^{r}_{rol}$	2.97	3.40	3.05	3.44	2.45	2.60
$PSY15^{i}$	**1.00**	1.45	**1.09**	1.43	1.18	1.31
$PSY15^{r}$	0.71	**1.24**	0.80	**1.41**	0.86	1.08
$\mathcal{B}^{Comb}_{\tau,1}$	5.20	4.26	5.16	4.34	5.56	5.44
$\mathcal{B}^{Comb}_{\tau,2}$	3.28	3.09	3.21	2.87	3.61	3.63
$\mathcal{B}^{Comb}_{\tau,3}$	1.95	2.06	2.02	2.17	1.92	2.01
$\mathcal{B}^{Comb}_{\tau,4}$	**1.16**	1.64	1.27	1.70	1.24	1.37
$\mathcal{B}^{Comb}_{\tau,5}$	0.73	1.54	**0.75**	1.51	**0.84**	**0.95**
$\mathcal{B}^{Comb}_{\tau,6}$	0.43	**1.20**	0.44	**1.23**	0.50	0.65
$\mathcal{B}^{Comb}_{\tau,7}$	0.21	0.76	0.18	0.73	0.29	0.32
$\mathcal{B}^{Comb}_{\tau,8}$	0.08	0.30	0.04	0.21	0.12	0.08
$\mathcal{B}^{RW}_{\tau}$	1.18	1.52	1.26	1.52	1.45	1.49

Notes: The table shows the average number of detected bubbles in the one-bubble simulation exercise. Rows 2-4 indicate the sample size, the emergence date and duration of the bubble. Rows 5-12 (middle panel) display the results for the individual indicators, rows 13-21 (lower panel) for the combination indicators. Bold entries indicate the best indicator in each indicator class (individual, combined).

Table 2.C.3: Frequency of detecting zero, one, two, or more bubbles

	Number of detected bubbles							
	Zero	**One**	**Two**	**More**	**Zero**	**One**	**Two**	**More**
Emergence date	$\tau_e = 40$				$\tau_e = 120$			
HP_{rec}	0.021	0.221	0.319	0.439	0.004	0.249	0.331	0.417
HP_{rol}	0.006	0.085	0.348	0.560	0.001	0.216	0.328	0.455
$PWY11^{i}_{rec}$	0.216	**0.472**	0.231	0.080	0.259	0.403	0.211	0.126
$PWY11^{r}_{rec}$	0.342	0.372	0.192	0.093	0.338	0.354	0.180	0.129
$PWY11^{i}_{rol}$	0.026	0.125	0.234	0.615	0.024	0.130	0.229	0.617
$PWY11^{r}_{rol}$	0.042	0.134	0.239	0.585	0.026	0.130	0.237	0.607
$PSY15^{i}$	0.318	0.443	0.175	0.065	0.229	**0.520**	0.194	0.058
$PSY15^{r}$	0.517	0.323	0.112	0.049	0.453	0.359	0.140	0.049
$\mathcal{B}^{Comb}_{\tau,1}$	0.000	0.001	0.031	0.968	0.000	0.005	0.041	0.954
$\mathcal{B}^{Comb}_{\tau,2}$	0.002	0.098	0.234	0.666	0.001	0.118	0.245	0.636
$\mathcal{B}^{Comb}_{\tau,3}$	0.052	0.369	0.307	0.272	0.036	0.365	0.308	0.291
$\mathcal{B}^{Comb}_{\tau,4}$	0.208	**0.516**	0.198	0.077	0.172	0.502	0.235	0.091
$\mathcal{B}^{Comb}_{\tau,5}$	0.422	0.447	0.112	0.019	0.406	0.461	0.114	0.019
$\mathcal{B}^{Comb}_{\tau,6}$	0.655	0.269	0.065	0.011	0.641	0.290	0.063	0.006
$\mathcal{B}^{Comb}_{\tau,7}$	0.819	0.150	0.029	0.002	0.846	0.131	0.021	0.002
$\mathcal{B}^{Comb}_{\tau,8}$	0.928	0.064	0.007	0.000	0.962	0.034	0.004	0.001
$\mathcal{B}^{RW}_{\tau}$	0.239	0.468	0.199	0.094	0.127	**0.591**	0.207	0.076

Notes: The table shows the frequency with which an indicator detects zero, one, two or more bubbles in the one-bubble simulation exercise with sample size $T = 200$, duration $\tau_d = 20$ and either an early (left panel) or a late bubble start (right panel). Row 2 indicates the emergence date. Rows 3-10 display the results for the individual indicators, rows 11-18 for the combination indicators. Bold entries indicate the best indicator in each category.

Table 2.C.4: Share of false decisions (Type I and II errors)

Indicator	**Share of false decisions**					
T	200	200	200	200	300	300
τ_e	40	40	120	120	60	180
τ_d	20	40	20	40	30	30
HP_{rec}	0.172	0.134	0.149	0.131	0.174	0.147
HP_{rol}	0.183	0.228	0.145	0.128	0.184	0.178
$PWY11^{i}_{rec}$	**0.081**	**0.106**	0.109	0.144	**0.078**	0.098
$PWY11^{r}_{rec}$	0.104	0.162	0.126	0.199	0.099	0.114
$PWY11^{i}_{rol}$	0.122	0.170	0.118	0.165	0.098	0.087
$PWY11^{r}_{rol}$	0.144	0.195	0.138	0.180	0.120	0.108
$PSY15^{i}$	0.085	0.118	**0.080**	**0.105**	0.081	**0.070**
$PSY15^{r}$	0.102	0.177	0.100	0.162	0.101	0.095
$\mathcal{B}^{Comb}_{\tau,1}$	0.260	0.265	0.223	0.181	0.246	0.218
$\mathcal{B}^{Comb}_{\tau,2}$	0.149	0.124	0.150	0.135	0.122	0.128
$\mathcal{B}^{Comb}_{\tau,3}$	0.088	**0.093**	0.087	0.094	0.076	0.070
$\mathcal{B}^{Comb}_{\tau,4}$	**0.079**	0.102	0.081	0.101	**0.074**	0.069
$\mathcal{B}^{Comb}_{\tau,5}$	0.083	0.127	0.085	0.126	0.080	0.078
$\mathcal{B}^{Comb}_{\tau,6}$	0.092	0.158	0.092	0.155	0.090	0.088
$\mathcal{B}^{Comb}_{\tau,7}$	0.096	0.179	0.097	0.179	0.095	0.095
$\mathcal{B}^{Comb}_{\tau,8}$	0.099	0.193	0.099	0.195	0.098	0.099
$\mathcal{B}^{RW}_{\tau}$	0.081	0.100	**0.070**	**0.085**	0.076	**0.060**

Notes: The table shows the number of false decisions (Type I and Type II errors) as a share of the sample length. For further notes, see Table 2.C.2.

Table 2.C.5: Probability of false signals (Type II errors)

Indicator	**Share of false signals over all signals**					
T	200	200	200	200	300	300
τ_e	40	40	120	120	60	180
τ_d	20	40	20	40	30	30
HP_{rec}	0.577	0.289	0.536	0.298	0.595	0.531
HP_{rol}	0.673	0.528	0.536	0.302	0.681	0.641
$PWY11^{i}_{rec}$	0.411	**0.169**	0.662	0.320	**0.321**	0.536
$PWY11^{r}_{rec}$	0.676	0.378	0.896	0.638	0.603	0.838
$PWY11^{i}_{rol}$	0.631	0.386	0.599	0.371	0.507	0.427
$PWY11^{r}_{rol}$	0.769	0.484	0.716	0.432	0.713	0.589
$PSY15^{i}$	**0.403**	0.175	**0.379**	**0.171**	0.356	**0.296**
$PSY15^{r}$	0.635	0.328	0.607	0.301	0.622	0.529
$\mathcal{B}^{Comb}_{\tau,3}$	0.424	0.158	0.414	0.179	0.342	0.303
$\mathcal{B}^{Comb}_{\tau,4}$	0.326	0.108	0.346	0.133	0.246	0.240
$\mathcal{B}^{Comb}_{\tau,5}$	**0.232**	**0.073**	0.313	**0.121**	**0.173**	**0.212**
$\mathcal{B}^{Comb}_{\tau,6}$	0.235	0.086	**0.297**	0.124	0.188	0.215
$\mathcal{B}^{RW}_{\tau}$	0.370	0.128	0.309	0.128	0.323	0.246

Notes: The table shows the number of false signals (Type II errors: signal if no bubble is present) as a share of all signals issued by the respective indicator. For further notes, see Table 2.C.2.

Table 2.C.6: Bubble probability if no signal is issued (Type I errors)

Indicator	**Probability of bubble if no alarm**					
T	200	200	200	200	300	300
τ_e	40	40	120	120	60	180
τ_d	20	40	20	40	30	30
HP_{rec}	*0.056*	*0.064*	*0.047*	*0.053*	*0.046*	*0.038*
HP_{rol}	*0.056*	*0.071*	*0.047*	*0.054*	*0.046*	*0.038*
$PWY11^{i}_{rec}$	**0.068**	**0.100**	0.084	0.126	**0.069**	0.083
$PWY11^{r}_{rec}$	0.087	0.154	0.097	0.177	0.090	0.097
$PWY11^{i}_{rol}$	0.069	0.139	*0.066*	0.135	*0.067*	*0.057*
$PWY11^{r}_{rol}$	0.086	0.158	0.081	0.146	0.088	0.077
$PSY15^{i}$	0.077	0.116	**0.071**	**0.103**	0.075	**0.063**
$PSY15^{r}$	0.094	0.175	0.091	0.159	0.094	0.088
$\mathcal{B}^{Comb}_{\tau,3}$	**0.062**	**0.080**	**0.058**	**0.075**	**0.060**	**0.051**
$\mathcal{B}^{Comb}_{\tau,4}$	0.071	0.102	0.070	0.098	0.069	0.063
$\mathcal{B}^{Comb}_{\tau,5}$	0.081	0.131	0.082	0.128	0.079	0.076
$\mathcal{B}^{Comb}_{\tau,6}$	0.091	0.162	0.091	0.158	0.090	0.087
$\mathcal{B}^{RW}_{\tau}$	0.071	0.097	0.060	0.081	0.067	**0.051**

Notes: The table shows the number of missed bubble periods (Type I errors: no signal if bubble is present) as a share of all periods in which the respective indicators does not issue a signal. The best indicator in each class that also achieves an overall low error rate (see Table 2.C.4) is highlighted in **bold**. The overall lowest Type I error rate is highlighted in *italics*.

Table 2.C.7: Signal instability

Indicator	**Signal instability**					
T	200	200	200	200	300	300
τ_e	40	40	120	120	60	180
τ_d	20	40	20	40	30	30
HP_{rec}	*0.066*	*0.061*	0.021	*0.009*	*0.077*	*0.011*
HP_{rol}	*0.066*	*0.061*	*0.018*	0.018	*0.077*	0.025
$PWY11^{i}_{rec}$	**0.129**	**0.163**	**0.023**	**0.027**	**0.108**	**0.020**
$PWY11^{r}_{rec}$	0.241	0.476	0.148	0.191	0.281	0.154
$PWY11^{i}_{rol}$	0.268	0.653	0.298	0.672	0.282	0.247
$PWY11^{r}_{rol}$	0.310	0.836	0.339	0.746	0.378	0.332
$PSY15^{i}$	0.156	0.215	0.094	0.125	0.197	0.104
$PSY15^{r}$	0.234	0.572	0.233	0.470	0.321	0.299
$\mathcal{B}^{Comb}_{\tau,3}$	0.155	0.209	0.102	0.162	**0.156**	0.110
$\mathcal{B}^{Comb}_{\tau,4}$	0.160	0.278	0.130	0.239	0.179	0.133
$\mathcal{B}^{Comb}_{\tau,5}$	0.176	0.376	0.152	0.309	0.214	0.158
$\mathcal{B}^{Comb}_{\tau,6}$	0.206	0.523	0.173	0.478	0.289	0.229
$\mathcal{B}^{RW}_{\tau}$	**0.122**	**0.122**	**0.093**	**0.117**	0.158	**0.091**

Notes: The table shows the average number of missing bubble signals between the first emergence and last collapse signal during the course of the true bubble. For further notes, see Table 2.C.6.

Table 2.C.8: Average signal delay

Indicator	**Emergence**						**Collapse**					
T	200	200	200	200	300	300	200	200	200	200	300	300
τ_e	40	40	120	120	60	180	40	40	120	120	60	180
τ_d	20	40	20	40	30	30	20	40	20	40	30	30
HP_{rec}	*8.19*	*9.41*	*8.05*	*8.79*	*10.20*	*9.91*	0.83	0.88	1.45	1.02	0.96	1.12
HP_{rol}	*8.19*	*9.41*	*7.87*	*8.53*	*10.20*	*9.01*	0.80	0.87	1.06	0.94	0.88	0.87
$PWY11^{i}_{rec}$	**8.60**	**13.65**	11.92	19.56	**14.47**	18.52	-0.30	-0.55	0.56	0.45	-0.13	0.69
$PWY11^{r}_{rec}$	7.75	16.70	10.29	23.35	13.69	17.27	-3.59	-4.35	-1.34	-1.33	-4.10	-1.92
$PWY11^{i}_{rol}$	8.00	10.82	*7.96*	*10.22*	*13.52*	*11.95*	-1.50	-7.48	-1.96	-7.98	-0.97	-1.11
$PWY11^{r}_{rol}$	8.47	13.92	9.14	12.96	15.16	14.81	-3.22	-7.39	-2.85	-6.94	-3.75	-2.50
$PSY15^{i}$	10.37	15.87	**10.30**	**14.89**	15.45	**14.08**	**-0.05**	**-0.03**	**0.29**	**0.28**	**-0.12**	**0.41**
$PSY15^{r}$	10.55	20.17	11.61	19.95	16.82	17.16	-3.70	-6.41	-1.95	-3.04	-5.25	-2.11
$\mathcal{B}^{Comb}_{\tau,3}$	**9.16**	**11.59**	9.40	**11.16**	**13.93**	12.78	0.28	0.38	0.34	0.27	0.30	0.51
$\mathcal{B}^{Comb}_{\tau,4}$	10.29	14.24	10.86	14.18	15.62	15.02	**-0.09**	**-0.22**	0.06	-0.21	**0.17**	0.38
$\mathcal{B}^{Comb}_{\tau,5}$	11.32	16.84	12.13	17.58	17.17	17.40	-0.43	-1.73	**-0.06**	-0.92	-0.27	**0.12**
$\mathcal{B}^{Comb}_{\tau,6}$	12.32	20.23	13.11	20.49	18.91	19.10	-1.03	-3.95	-0.57	-2.38	-1.14	-0.49
$\mathcal{B}^{RW}_{\tau}$	10.22	14.51	**9.18**	12.13	14.84	**12.00**	0.41	0.45	0.50	0.38	0.31	0.43

Notes: The table shows the average delay before an indicator signals the emergence or collapse of the true bubble. For further notes, see Table 2.C.6.

Table 2.C.9: Bubble detection frequency and accuracy: Two-bubble scenario

		Detection rates				Type I + II errors		
Indicator	**Average**	**Zero**	**One**	**Two**	**More**	**I + II**	**II**	**I**
HP_{rec}	2.80	0.000	0.041	*0.426*	0.533	0.174	0.375	*0.114*
HP_{rol}	2.87	0.000	0.029	0.394	0.577	*0.161*	0.359	*0.103*
$PWY11^{i}_{rec}$	1.25	0.191	0.475	0.254	0.080	0.184	0.458	0.174
$PWY11^{r}_{rec}$	1.04	0.340	0.385	0.193	0.082	0.202	0.666	0.190
$PWY11^{i}_{rol}$	3.29	0.017	0.083	0.216	0.685	0.172	0.396	*0.144*
$PWY11^{r}_{rol}$	2.95	0.033	0.132	0.245	0.590	0.201	0.530	0.170
$PSY15^{i}$	**1.57**	0.170	0.312	**0.346**	0.171	**0.167**	**0.344**	**0.160**
$PSY15^{r}$	0.93	0.403	0.356	0.172	0.068	0.194	0.475	0.189
$\mathcal{B}^{Comb}_{\tau,1}$	4.66	0.000	0.001	0.047	0.952	0.190	0.450	*0.082*
$\mathcal{B}^{Comb}_{\tau,2}$	3.22	0.000	0.011	0.321	0.668	0.150	0.333	*0.097*
$\mathcal{B}^{Comb}_{\tau,3}$	2.39	0.014	0.125	**0.479**	0.381	**0.148**	0.283	**0.134**
$\mathcal{B}^{Comb}_{\tau,4}$	1.84	0.084	0.284	0.406	0.226	0.161	**0.281**	0.155
$\mathcal{B}^{Comb}_{\tau,5}$	1.20	0.266	0.391	0.241	0.101	0.178	0.300	0.176
$\mathcal{B}^{Comb}_{\tau,6}$	0.65	0.535	0.316	0.121	0.028	0.191	0.323	0.189
$\mathcal{B}^{Comb}_{\tau,7}$	0.25	0.790	0.172	0.034	0.003	0.197	0.336	0.196
$\mathcal{B}^{Comb}_{\tau,8}$	0.09	0.923	0.067	0.010	0.001	0.199	0.313	0.199
$\mathcal{B}^{RW}_{\tau}$	**1.85**	0.087	0.245	0.451	0.217	0.153	0.298	0.145

Notes: The table shows the average number of detected bubbles (first column), the frequency distribution of detection zero, one, two, or more bubbles (second to fifth column) as well as the overall error rate (column six, as in Table 2.C.4), the share of false signals over all signals (column seven, as in Table 2.C.5), and the probability of missing a bubble if no signal is issued (column eight, as in Table 2.C.6). See Table 2.C.6 for further notes.

Table 2.C.10: Predictive accuracy: Stock price bubbles

	Forecast horizon (in months)							
Model	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**
Benchmark	0.007	0.012	0.022	0.035	0.044	0.054	0.078	0.101
HP_{rec}	0.988*	0.980*	0.969*	0.986*	1.008*	1.008*	0.990	0.963*
HP_{rol}	0.994*	0.984*	0.973*	0.980*	1.002	1.016*	1.030*	1.051*
$PWY11^{i}_{rec}$	0.995*	0.986*	0.970*	0.976*	0.996*	1.003*	0.993	0.959*
$PWY11^{r}_{rec}$	0.999	0.994*	0.988*	1.003	1.018*	1.024*	1.036	1.047
$PWY11^{i}_{rol}$	**0.979***	**0.959***	**0.942***	**0.955***	0.981*	0.996	1.001	0.996
$PWY11^{r}_{rol}$	1.021	1.032	1.042	1.056*	1.057*	1.041*	1.012*	1.000
$PSY15^{i}$	0.994*	0.988*	0.976*	0.974*	**0.977***	**0.972***	**0.952***	**0.924***
$PSY15^{r}$	1.024	1.036	1.041	1.046	1.043*	1.030*	0.999	0.979
$\mathcal{B}^{Comb}_{\tau,1}$	1.011	1.018	1.024*	1.028*	1.030*	1.026*	1.007*	1.010*
$\mathcal{B}^{Comb}_{\tau,2}$	1.008	1.011	1.010	1.021	1.033*	1.030*	0.993*	0.961*
$\mathcal{B}^{Comb}_{\tau,3}$	**0.970***	**0.954***	**0.929***	**0.940***	**0.967***	0.983*	0.979*	0.948*
$\mathcal{B}^{Comb}_{\tau,4}$	0.972*	0.954*	0.930*	0.944*	0.969*	0.983*	0.982*	0.959*
$\mathcal{B}^{Comb}_{\tau,5}$	0.974*	0.957*	0.934*	0.951*	0.978*	0.991*	0.996	0.986
$\mathcal{B}^{Comb}_{\tau,6}$	0.992*	0.976*	0.961*	0.970*	0.983*	**0.982***	**0.966***	**0.945***
$\mathcal{B}^{Comb}_{\tau,7}$	0.992*	0.980*	0.975*	0.984*	0.997	0.995	0.986	0.969
$\mathcal{B}^{Comb}_{\tau,8}$	0.996*	0.989*	0.981*	0.984*	0.994*	0.994	0.990	0.982
$\mathcal{B}^{RW}_{\tau}$	0.987*	0.973*	0.957*	0.969*	0.993*	1.004*	0.997	0.969*

Notes: Rows two and following display the RMSPE of the stock price bubble augmented models relative to the RMSPE of the benchmark given in the first row. Values less than 1 indicate that the augmented model is superior to the benchmark. Forecasts denoted by * are significantly different at the 5%-level from the benchmark as indicated by the test for equal predictive ability for real-time data by Clark and McCracken (2009). Bold values show the best forecast for a given horizon in each model class. The CPI, IP and the unemployment rate are published with a lag of one month. The past five months of IP are subject to revisions, while the seasonal factors in CPI and the unemployment rate are recalculated for up to five years.

CHAPTER 3

Predicting Output with Real-Time Bubble Indicators[1]

3.1 Introduction

Accurate forecasts of future economic growth are of paramount importance for the decision-making of economic policymakers. The Great Recession of 2008-09, however, has highlighted the challenges that forecasters and economic observers face in this task. Yet, the recession has also emphasized that a stronger surveillance of asset market conditions may provide valuable information for predicting future economic activity. In particular, the recession has forcefully illustrated the impact of pronounced asset price movements on the real economy. Following the collapse of the housing bubble in 2005-06 the U.S. economy in December 2007 entered the longest and deepest recessionary period since the Great Depression of 1929. Similarly, the crash of the dot-com bubble in March 2000 ended the decade-long expansion of the U.S. economy, with the recession starting in March 2001.[2]

This suggests that bubbles on stock and housing markets, if detected in real-time, could be a promising predictor for future economic activity. While this real-time detectability of asset price bubbles has, however, long been contested (cf. Trichet, 2005; Kohn, 2006), Phillips et al. (2011, 2015) have recently challenged this conventional view. In particular, the authors propose a recursive monitoring approach that

[1]This chapter is based on joint work with Dirk Ulbricht. We are thankful to Helmut Lütkepohl and Konstantin A. Kholodilin for their helpful comments and advice.

[2]See "US Business Cycle Expansions and Contractions", The National Bureau for Economic Research (NBER), available online: `www.nber.org/cycles.html`

is capable of detecting periods that display patterns typical for asset price bubbles by testing both an asset's price and its fundamental series for explosive roots at each point in the sample. By this, the authors have provided a new, promising approach to a literature that has previously used price deviations from an (HP-)filtered trend as a proxy for excesses on asset markets (e.g. Detken and Smets, 2004; Adalid and Detken, 2007; Assenmacher-Wesche and Gerlach, 2010).

In this paper, we hence exploit the real-time information from such binary bubble indicators to forecast a popular monthly index for economic activity, namely total industrial production (IPT). In particular, we assess the predictive content of several specifications of the bubble indicators of Phillips et al. (2011, 2015), of the popular indicators based on the HP-filter, and of combinations of all such indicators as developed in Chapter 2. Against the background of the pronounced stock and house price bubbles the U.S. experienced in the past two decades we apply these indicators to both U.S. stock and housing market data.[3]

If these indicators are indeed capable of accurately detecting asset price bubbles in real-time, there is strong reason to believe that this information may help to predict economic activity both at short and longer forecast horizons. At the longer horizons, this predictive content may result from the ensuing recession caused by a future burst of the bubble. In the short run, in contrast, the bubble is likely to persist and thus to contribute to higher economic growth due to financial accelerator mechanism. Here, Kiyotaki and Moore (1997) and Bernanke et al. (1999) note that rising asset prices increase the net worth of firms and households, thereby raising the value of their collateral and their borrowing limits; and lowering their default probability and the external finance premium demanded by lenders. Similarly, rising asset prices also increase the equity position of lenders, allowing them to increase their supply of credit (Holmstrom and Tirole, 1997). All else equal, an increase in asset prices thus stimulates investment, consumption, and, ultimately, output. Importantly, these financial accelerator mechanisms may be particularly pronounced during speculative asset price bubbles when optimism about future capital gains reinforces current capital gains and induces a broad surge in lending and investment

[3]The real economic impacts are likely to be largest for debt-financed asset bubbles (Borio and Lowe, 2004; Schularick and Taylor, 2012). A natural extension of the present paper is thus to investigate the predictive content of excessive credit growth. This is not done here, as credit data is only available on the quarterly frequency, limiting the scope for forecast evaluations.

(cf. Bernanke and Gertler, 1999; Martin and Ventura, 2011, 2012). Thus, the signal of an asset price bubble is likely to carry different implications for output growth over different future horizons depending on when the bubble bursts.

We therefore assess the value of the available real-time asset price bubble indicators in predicting output growth for up to two years in a monthly sample from 1975 to 2014. Here, we generally follow the recursive out-of-sample forecast experiment of Stock and Watson (2003), evaluating the predictive gain of these indicators against a large set of macroeconomic and financial predictors commonly used to forecast output. Further, we assess if these bubble indicators can also improve upon an autoregressive, dynamic factor model (DFM) that condenses the information from the set of macroeconomic and financial predictors at each point in the sample. Throughout the paper, we acknowledge the real-time availability of all data, taking into account all publication lags and revisions.

We find that several indicators for asset price bubbles strongly improve upon the autoregressive benchmark model for output and are routinely included in the Model Confidence Set (MCS) of superior models. In particular, the best stock price bubble indicators are among the best ten of in total 248 forecast models for forecast horizons from three to 18 months. In addition, we note that these indicators also provide the most plausible bubble signals. Since bubble episodes on stock and housing markets are not uniquely defined, and the available indicators provide differing bubble signals, our forecast assessment hence also offers an additional criterion to address the quality of the indicators' real-time bubble signals. House price bubbles, however, are more difficult to detect but provide significant predictive gains at the two year horizon. We also show that stock and house price bubble indicators provide predictive gains to the DFM that condenses the information of all rival macroeconomic and financial predictors.

Finally, we address the question at which times asset price bubble indicators are particularly useful for forecasting output. We again find that the most promising bubble indicators outperform the simple autoregressive forecasts both during recessions and expansions. Furthermore, the relative gains to the AR model are quantitatively comparable across the two regimes. This suggests that these bubble indicators add valuable information to lagged values of output growth independent of time. In contrast, the predictive accuracy of the conventional, non-bubble in-

dicators varies strongly over time as noted also by Stock and Watson (2003) and Rossi and Sekhposyan (2010) with their predictive gains stemming largely from the three short recessionary periods in our sample. For expansion periods the predictive value of conventional macroeconomic and financial indicators is considerably smaller. In contrast, adding the binary indicators for asset price bubbles to the AR model strongly improves its forecast. Specifically, the eight best forecasts during expansions are all provided by bubble-augmented AR models. All in all, we thus show that the most accurate real-time indicators for asset price bubbles are promising alternative predictors for real economic activity. By this, we provide evidence that these indicators are consistently linked to amplifying or dampening output at different horizons as predicted by the theoretical literature. This suggests that these indicators are indeed capable of detecting true bubble episodes in real-time.

The remainder of this paper is structured as follows. The next section introduces the bubble indicators and applies them to U.S. stock and housing market data. Further, we also present the set of alternative predictors and their transformations in this section. Section 3.3 then describes the forecast experiment and the evaluation statistics. Section 3.4 presents the results, and Section 3.5 concludes.

3.2 Real-time indicators for asset price bubbles

We assess the predictive content of three classes of real-time monitoring indicators for asset price bubbles: six specifications of the recursive (or rolling) tests for explosive roots in asset prices and fundamentals developed by Phillips et al. (2011, 2015), two specifications of indicators based on price deviations from one-sided HP-filtered trends, and nine specifications of combination indicators developed in Chapter 2. This provides us with a total of 17 real-time monitoring indicators for asset price bubbles. In the following, we briefly describe all indicators which follow exactly the specifications outlined in Chapter 2. The reader of the previous chapter may hence skip to Section 3.2.4.

3.2.1 Detecting explosive bubble processes by unit-root tests

The tests of Phillips et al. (2011, 2015) build on the insights of Diba and Grossman (1988) that a rational asset price bubble can only be sustained when the bubble

component in an asset's price grows exponentially at the rate of interest in expectation.[4] Assuming that the underlying fundamental series (e.g. dividends D_t) is $I(1)$, this implies that any explosive behavior of the observable price series (P_t) must stem from the presence of a (rational) bubble.[5] The monitoring approaches by Phillips et al. (2011, 2015) hence build on sequentially testing both the price and the fundamental series for explosive growth.

Specifically, these approaches apply right-tailed ADF tests of the null hypothesis $H_0 : \delta = 1$ against its alternative $H_1 : \delta > 1$ by estimating the ADF equation

$$z_t = \mu_z + \delta z_{t-1} + \sum_{j=1}^{J} \phi_j \Delta z_{t-j} + v_t, \quad t = 1, \ldots, \tau, \quad v_t \overset{iid}{\sim} N(0, \sigma_v^2) \tag{3.1}$$

for both prices (p_t) and fundamentals (d_t) individually, i.e. $z_t \in \{p_t, d_t\}$, or for their log-ratio, i.e. $z_t = p_t - d_t$, with $p_t = \log(P_t)$ and $d_t = \log(D_t)$. Specifically, these tests are carried out sequentially at each margin in the sample, i.e. for each $\tau = \tau_0, \tau_0+1, \ldots, T$. This can be done by recursive or rolling regressions of (3.1) and provides a series of ADF tests for prices ($ADF_{p,\tau}$) and fundamentals ($ADF_{d,\tau}$) when both series are tested individually. A bubble is then signaled at each τ for which $ADF_{p,\tau} \geq cv_{\alpha_T}^{adf}(\tau)$, while $ADF_{d,\tau} < cv_{\alpha_T}^{adf}(\tau)$. Similarly, when testing the log-ratio, a bubble is signaled at each τ for which $ADF_{pd,\tau} \geq cv_{\alpha_T}^{adf}(\tau)$.

As emphasized in Chapter 2, the choice of whether testing prices and fundamentals sequentially, or testing their ratio, influences the outcome. In particular, under the null hypothesis of no bubble the log-ratio is stationary (Campbell and Shiller, 1987, 1988; Cochrane, 1992). Testing the individual series thus provides a larger power against the null hypothesis than testing the log-ratio. While this implies a higher chance of detecting explosive growth in prices in the bubble scenario, it also raises the chance of false positive signals about explosive growth in the dividend series. In this case, a true bubble would be missed. In contrast, testing the log-ratio assesses the relative growth rates of the two series, allowing to detect periods of an emerging asset price overvaluation when prices grow faster than dividends.

[4] See Camerer (1989) for a derivation of this condition and further extensions of the rational bubble framework.

[5] Scherbina and Schlusche (2014) provide a survey on behavioral explanations for explosive, bubbly asset price increases.

First, we employ the binary bubble indicators of Phillips et al. (2011) by estimating (3.1) either recursively (denoted by $PWY11^{\bullet}_{rec}$) or by rolling windows ($PWY11^{\bullet}_{rol}$). Here, $\bullet \in \{i, r\}$ denotes the indicator based on testing the individual series ($\bullet = i$), or based on testing their ratio ($\bullet = r$). These indicators take the value one for each τ for which a bubble is signaled. The recursive approach starts at $\tau_0 = \lfloor r_0 T \rfloor$, where $r_0 \in (0, 1)$ is the minimum share of T for which (3.1) is estimated. For the rolling estimation approach, the window size is fixed at τ_0. Critical values are given by $cv^{adf}_{\alpha_T}(\tau) = \log(\log(\tau))/100$ for the recursive approach, and a constant $cv^{adf}_{\alpha_T}(\tau) = \log(log(\tau_0))/100$ for the rolling window estimation (Phillips et al., 2011).[6]

While Phillips et al. (2011) show that their approach is capable of detecting single bubble processes in asset prices, Phillips et al. (2015) find that it has troubles to "restart" after a first bubble has collapsed, and frequently misses any second or further bubble in the sample. Therefore, the authors extend the work of Phillips et al. (2011) and augment the forward recursive regressions by testing all possible backward extending windows at each given margin τ_2. Specifically, this approach not only moves the estimation end point τ_2 forward for each recursive regression from $\tau_2 = \tau_0, \ldots, T$, but also assesses all possible backward expanding windows from the start point $\tau_1 \in [0, \max(\tau_2 - \tau_0, 1)]$ for a given τ_2. For a given end point $\tau_2 \in [\tau_0, T]$ and all possible start points $\tau_1 \in [1, \max(\tau_2 - \tau_0, 1)]$, the sequence of ADF test statistics is denoted by $\{ADF^{\tau_2}_{\tau_1}\}$. The unique test statistic at τ_2, $BSADF_{\tau_2}$, is then obtained by taking the supremum of this sequence, i.e. $BSADF_{\tau_2} = \sup_{\tau_1 \in [1, \tau_2 - \tau_0]} \left\{ADF^{\tau_2}_{\tau_1}\right\}$. The two binary indicators $PSY15^{\bullet}$, with $\bullet = \{i, r\}$, then take the value one for each τ_2 for which $BSADF_{\tau_2}$ exceeds the critical value $cv^{bsadf}_{\alpha_T}(\tau_2)$.[7]

[6] Following the authors, we set $r_0 = 0.1$ for the recursive estimation and $r_0 = 0.2$ for the rolling window approach. The lag order J is determined at each τ by the Akaike Information Criterion (AIC) with $J^{max} = 12$.

[7] Critical values for the $BSADF_{\tau_2}$ statistic are obtained from simulations using the MATLAB code provided by Shu-Ping Shi on her website (`https://sites.google.com/site/shupingshi/home/research`). Following Phillips et al. (2015), the initial sample size is set to $r_0 = 0.01 + \frac{1.8}{\sqrt{T}}$ and the lag order in (3.1) is fixed to $J = 1$.

3.2.2 Bubbles as price deviations from an HP-trend

Prior to Phillips et al. (2011), the literature commonly defined asset price bubbles as an "excessive" deviation of the real price series from its one-sided HP-filtered trend. This builds on the assumption that the asset's fundamentals follow a slow-moving trend (e.g. Detken and Smets, 2004; Adalid and Detken, 2007; Assenmacher-Wesche and Gerlach, 2010). As such, this indicator can potentially capture periods in which large deviations of prices from their past trend occur, regardless of the speed of this accumulation process. We therefore also include two specifications of bubble indicators based on the HP-filter. Following Assenmacher-Wesche and Gerlach (2010), we define a bubble to occur when the real asset price exceeds its trend by a threshold of at least κ_{hp}, with $\kappa^s_{hp} = 10\%$ for stock price bubbles and $\kappa^h_{hp} = 7.5\%$ for house price bubbles. The smoothing parameter λ for estimating the trend component follows Assenmacher-Wesche and Gerlach (2010) but is adjusted to the monthly frequency, i.e. $\lambda = 8{,}100{,}000$. In addition to a recursive trend estimate, denoted by HP_{rec}, we also employ a rolling estimation with window size $\omega = 96$, denoted by HP_{rol}. Following the literature, we do not update past estimates of the trend as new observations arrive, but use the margin τ and past price observation only to estimate the τ trend value.

3.2.3 Combination approaches to real-time bubble detection

As shown by Phillips et al. (2015) and extended in Chapter 2 all individual indicators feature complementary strengths and weaknesses depending on their exact specification, and depending on the characteristics of the bubble process. For instance, Phillips et al. (2015) show that their indicator more often detects a second bubble in the sample compared to the indicators of Phillips et al. (2011), which, however, comes at the cost of a lower power to detect the first bubble. A natural way to aggregate the information content across indicators is thus to combine signals from all individual indicators. This can be implemented following the two approaches developed in Chapter 2. As shown in this chapter several of these combinations outperform the individual indicators with regard to their signal accuracy. We hence also assess the predictive ability of these combination indicators.

First, we aggregate the information content of the individual indicators based on the simple counting approach developed in Chapter 2. In particular, this indicator

counts how many of the eight individual indicators $\mathfrak{B} = \{PWY11^i_{rec}, PWY11^r_{rec}, PWY11^i_{rol}, PWY11^r_{rol}, PSY15^i, PSY15^r, HP_{rec}, HP_{rol}\}$ simultaneously signal a bubble. If this number exceeds a threshold κ, the combination indicator $\mathcal{B}^{Comb}_{\tau,\kappa}$ takes the value one. The caveat of this approach is that the optimal choice for κ is not clear. Therefore, we explore the predictive value for all possible choices for $\kappa = 1, \ldots, 8$. To rule out the trivial case of $\mathcal{B}^{Comb}_{\tau,\kappa} = 0 \; \forall \; \tau$, we require that $\mathcal{B}^{Comb}_{\tau,\kappa} = 1$ for at least one τ.[8]

Second, we employ the econometric combination indicator developed in Chapter 2. This indicator addresses the limitation of the indeterminacy of the threshold choice κ and allows to control the overall size of the approach. Specifically, this combination builds on the multiple testing approach of Romano and Wolf (2005), taking into account the correlation structure between the individual test statistics. For this, it is required that all included tests provide comparable test statistics which rules out the inclusion of the HP indicators. As in Chapter 2, we will use this algorithm to combine the three indicators $PWY11^i_{rec}$, $PWY11^i_{rol}$, and $PSY15^i$. The general idea of this test is to order the (studentized) test statistics associated with the three indicators at each margin τ by size, and then to assess whether the largest one exceeds the critical value. This critical value is obtained from bootstrap simulations and controls the familywise error rate (FWE) to be no greater than $\alpha = 0.05$, with the FWE being defined as the probability that at least one test falsely rejects the true null hypothesis $H_0 : \delta = 1$. We denote the combination indicator obtained from this application of the Romano and Wolf (2005) algorithm as $\mathcal{B}^{RW}_{\tau}$.

3.2.4 Stock and house price bubbles in the U.S.

Figures 3.A.1 to 3.A.4 display the signals obtained from all individual and combination indicators when applied to monthly U.S. stock and housing market data for the period from 1975M1 to 2014M12. For stock prices, we choose the S&P 500 index as the most relevant stock market index that captures around 75% of the total U.S. market capitalization. Naturally, the underlying fundamental series of the stock price index is the associated dividend series. For house prices, however, the choice of the underlying fundamental series is more controversial. Here, we follow Pavlidis

[8] This provides us with eight counting combination indicators $\mathcal{B}^{Comb}_{\tau,\kappa}$ for stock markets, while for house price data, $\mathcal{B}^{Comb}_{\tau,\kappa} = 0 \; \forall \; \tau$ for $\kappa > 5$.

et al. (2014) and employ real disposable income per capita as a measure for the fundamental determinant of house prices. By this we aim to measure the affordability of housing.[9] Assuming that households devote a constant share of their total income to renting, house prices should grow only at the growth rate of per capita real disposable income.[10]

Figure 3.A.1 shows the bubble episodes for the S&P 500 index as detected by the indicators of Phillips et al. (2011, 2015) and by the indicators based on the HP-filter. Overall, there is a strong heterogeneity in the provided signals, with only few indicators providing plausible signals throughout. The most plausible real-time bubble signals are provided by the $PWY11^{i}_{rol}$ and the $PSY15^{i}$ indicators, detecting both the bubble preceding the 1987 "Black Monday" crash and the dot-com bubble. Their counterparts testing the log-ratio of prices and dividends both signal a negative bubble following the global financial crisis. Both $PWY11^{\bullet}_{rec}$ indicators do not provide false positive signals after the crisis, but instead miss the crash of the dot-com bubble, providing signals well into 2002. Finally, the HP indicators capture the 1987 and dot-com bubbles at an early stage, yet generally provide the most (false) signals. Therefore, the $PWY11^{i}_{rol}$ and the $PSY15^{i}$ indicators can be expected to provide the best forecasts for output.

Figure 3.A.1 however also showed that some common periods for asset price bubbles emerge. Therefore, combinations of such indicators can be a promising tool to focus on these periods only, while discarding false positive signals by the different individual indicators. These signals from the combination indicators are displayed in

[9] All stock and house price data are obtained from the online supplement of Shiller (2005) available at `http://www.econ.yale.edu/~shiller/data.htm`. All price and dividend series are deflated by the real-time U.S. Consumer Price Index from the Real-Time Data Set of the Federal Reserve Bank of Philadelphia. The data for real disposable income per capita is obtained from FRED.

[10] This measure is of course a simplification as it discards changes in mortgage rates, or related tax rates. See Himmelberg et al. (2005) for a discussion of this issue. Nonetheless, we think that our measure is preferred to other approaches such as using rent series as the underlying fundamental for several reasons. First, rent series are generally measured with great error only and do not account for the intrinsic value of owning a house. Secondly, the causality structure between rental and purchase prices for housing is ambiguous. In case of high market power on the home owners' side, it is possible that rising purchase prices induce rising rent prices, thus leading to explosive growth in both series during a housing bubble. Eventually, this development is likely not sustainable, yet the indicators described above would not signal a bubble.

Figure 3.A.2.[11] For a threshold choice of $\kappa = 2$, the counting combination indicator appears to detect too many bubble episodes to be plausible. Choosing $\kappa = 6$, in contrast, appears to be too restrictive as this combination misses the 1987 bubble and provides an instable signal during the dot-com bubble. For threshold choices in between, however, the figure reveals the strengths of this approach. For $\kappa = 3, \ldots, 5$ only the pre-"Black Monday" bubble and the dot-com bubble are detected, with accurate collapse dates. Similarly, also the econometric combination $\mathcal{B}_{\tau}^{RW}$ provides reasonable signals for these two periods but the dot-com bubble collapse is only detected in 2001M2.

Figures 3.A.3 and 3.A.4 display the results for house price bubbles. Overall, the ratio of house prices to disposable income per capita is less volatile than the price to dividend ratio for the S&P 500 index. This appears to provide a strong challenge to the available bubble indicators, especially for the $PWY11^{\bullet}_{rec}$ indicators (failing to detect prolonged bubble episodes) and the $PSY15^{r}$ indicator (providing too many signals). The most plausible signals for the housing bubble preceding the global financial crisis are provided by the HP_{rec} and the $PWY11^{i}_{rol}$ indicators. When aggregating signals through combinations, again a threshold of $\kappa = 3$ or $\kappa = 4$ provides the most promising signals. Also, the econometric approach tends to provide bubble signals only during pronounced boom periods, yet its signal is inherently instable during the run of the housing bubble. Overall, the predictive ability of the most promising indicators for house price bubbles is likely to be limited due to the lack of a pronounced bubble early in the sample.

3.2.5 Other predictors

As a reference to compare the predictive ability of the $N = 31$ bubble indicators to other commonly used predictors for output growth we include a comprehensive set of 46 macroeconomic and financial variables. Specifically, we include real economic variables comprising measures of output, capacity utilization, the labor market, (energy and commodity) prices, as well as financial variables namely interest rates, their spreads, and exchange rates. We also include stock prices, dividends, earnings, house prices, and housing starts as related indicators to our bubble indicators. Further, in

[11] We suppress the signals from the simple counting indicators with $\kappa = 1, 7, 8$ for ease of presentation, noting that setting $\kappa = 1$ induces strong overdetection while combinations with $\kappa = 7, 8$ provide too few signals.

in our benchmark AR model specifications, we do not pre-test any of the predictors for stationarity but directly consider both levels as well as up to four transformations of each variable. These are first (D1) and second differences (D2), as well as month-on-month (P1) and year-on-year (P12) percentage changes.[12] This provides us with a total of $\widetilde{M} = 216$ candidate non-bubble predictors. We take into account the real-time availability of all data, considering both publication lags and revisions where necessary. Stock prices and dividends are unrevised and available immediately, such that the stock price bubble indicators are timely available. House prices and disposable income per capita are available with a publication lag of two months. For an overview of all variables, their publication lag, and the transformations see Table 3.B.2.

3.3 Real-time forecast experiment and evaluation

In this section, we first describe our pseudo-out-of-sample forecast experiment. Second, we outline our forecast evaluation statistics.

3.3.1 Model specifications

To evaluate the predictive content of the bubble indicators, we follow the set-up of Stock and Watson (1999, 2003), and assess each bubble indicator's forecast accuracy relative to an autoregressive (AR) benchmark model, and relative to the set of macroeconomic and financial predictors. Specifically, we first construct a large set of competitor models, consisting of the AR and the AR augmented by each of the exogenous predictor variables (ARX) individually. Second, we evaluate the value of the binary bubble indicators against a model that draws on all the information provided by the large set of macroeconomic and financial indicators. For this, we augment an autoregressive model by factors extracted from the set of predictor variables as in Stock and Watson (1999). We now describe these two forecast set-ups in order.

First, we compare the predictive ability of all candidate models to the benchmark univariate AR model of output growth. For this, we sequentially augment the AR

[12]Since spreads are zero at times, their percentage changes are not defined and thus discarded from the set of predictors. This is denoted by 'a' in the respective columns in Table 3.B.2.

model by one of the candidate predictors from Table 3.B.2 or from the set of stock and house price bubble indicators. Following the literature of, for example, Stock and Watson (1999, 2003), Rossi and Sekhposyan (2010), and Ulbricht et al. (2016), our forecast specification at margin τ for the individual model $m = 0, 1, \ldots, M$ ($M = 247$) for forecast horizon h is given by

$$y_{t+h}^h = \alpha_{m,h} + \sum_{p=l_0}^{P+l_0} \beta_{m,h,p} y_{t-p} + \sum_{q=l_m}^{Q_m+l_m} \gamma_{m,h,q} x_{m,t-q} + u_{m,t+h}^h, \quad t = 1, \ldots, \tau \qquad (3.2)$$

where y_{t+h}^h is the annualized growth rate of the IPT index at time t over the next h months, i.e. $y_{t+h}^h = \frac{1200}{h} \log\left(\frac{IPT_{t+h}}{IPT_t}\right)$. As in the literature, the lagged dependent variable is specified in first log-differences, i.e. $y_t = 1200 \log\left(\frac{IPT_t}{IPT_{t-1}}\right)$. Further, $x_{m,t}$ is the single candidate predictor for model m, $u_{m,t}^h$ the respective error term, and $\alpha_{m,h}, \beta_{m,h,p}, \gamma_{m,h,q}$ are (model-, horizon-, and lag-specific) regression coefficients. For the benchmark AR model ($m = 0$), all $\gamma_{m,h,q} = 0$ so that the third term disappears.[13] We take into account all dimensions of the real-time availability of all variables. For this, l_0 takes on the value of the publication lag of IPT (i.e. $l_0 = 1$) and l_m is the publication lag of the candidate predictor (as given in the fourth column of Table 3.B.2). Further, we consider the specific vintage τ data as available at the time of the forecast for all real economic predictors (see column six of Table 3.B.2). By this, we account for all revisions as well as recalculations of seasonal adjustment factors in real economic data as provided by the responsible statistical agency. Finally, the optimal lag lengths P and Q_m are determined at each forecast margin τ using the Bayesian Information Criterion (BIC) for a maximum lag length of twelve months. Specifically, we first determine the lag length P of the AR model, and holding this P fixed, we then estimate the optimal Q_m for each candidate model, but require $Q_m \geq 1$. We repeat this for each h-specific model at each margin τ.

Second, we evaluate the predictive value of the bubble indicators against an autoregressive dynamic factor model as employed by Stock and Watson (1999). The authors show that such models may provide better inflation forecasts than AR(X) models. Therefore, we follow the approach of Stock and Watson (1999) and construct factors as the principal components from the $\breve{M} = 180$ *stationary* macroeconomic

[13] If the candidate predictor shows no variation (e.g. no bubble signal) up to the forecast margin τ, the benchmark AR forecast is applied instead.

and financial predictors $\breve{X}_t = [x_{1,t}, \ldots, x_{\breve{M},t}]'$ from all $\widetilde{M} = 216$ transformations as available at time τ.[14] We thus obtain a k-dimensional vector of factors F_t with $0 < k << \breve{M}$ condensing the information from all conventional predictors. The optimal number of factors k is determined by the Bai and Ng (2002) criterion from a maximum of $K = 5$ factors.[15] We then adapt the autoregressive model in (3.2) by adding the factors as

$$y_{t+h}^h = \alpha_{i,h} + \sum_{p=l_0}^{P+l_0} \beta_{i,h,p} y_{t-p} + \sum_{r=1}^{R} \delta_{i,h,r} F'_{t-r} + \sum_{q=l_i}^{\tilde{Q}_i+l_i} \tilde{\gamma}_{i,h,q} b_{i,t-q} + u_{i,t+h}^h \tag{3.3}$$

where R is the optimal lag length of the vector of factors, and $b_{i,t-q}$ with $i = 1, \ldots, N$ is one indicator from the set of $N = 31$ stock and house price bubble indicators not included in the construction of factors. For this evaluation, the benchmark model is now the dynamic factor model (DFM) with all $\tilde{\gamma}_{i,h,q} = 0$.

However, constructing the factors F_t from the entire set of the $\breve{M}$ (transformed) macroeconomic and financial predictors can be criticized as some of these predictors may not provide valuable information for predicting output, but nonetheless assert strong influence on the shape of the factors. To discard these variables, we also construct factors from a preselected set of the most promising candidate predictors at each margin τ. For this, we follow Bai and Ng (2008) and employ the *Elastic Net* regularization and variable selection method of Zou and Hastie (2005). This method performs a penalized least squares estimation of the augmented AR

$$y_{t+h}^h = \alpha + \beta y_t + \gamma \breve{X}_t + e_{t+h}^h \tag{3.4}$$

where we again acknowledge the real-time availability of each predictor in $\breve{X}_t$. Specifically, the elastic net estimate of $\boldsymbol{\beta} = [\alpha, \beta, \gamma]'$ is given by $\hat{\boldsymbol{\beta}} = (1 + \lambda_2) \arg\min_{\boldsymbol{\beta}} |\boldsymbol{y} -$

[14] We define a predictor to be non-stationary if at least two of four tests indicate non-stationarity. The tests are the ADF- and the Phillips-Perron (PP) tests with a constant and a linear trend, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test with either level or trend stationarity as the null hypothesis (Perron, 1988; Kwiatkowski et al., 1992). We find all macroeconomic and financial predictors to be I(1) at most. The excluded level variables are denoted by 'y' in column 'L' of Table 3.B.2.

[15] Specifically, the criterion for determining the number of factors k at margin $\tau = 1, \ldots, T$ reads as $IC_{p2}(k) = log(V(k)) + k \frac{\breve{M}+\tau}{\breve{M}\tau} \log(\min(\breve{M}, \tau))$, where V(k) is the sum of squared residuals from the k-factor model.

$\boldsymbol{X}\boldsymbol{\beta}|^2 + \lambda_2|\boldsymbol{\beta}|^2 + \lambda_1|\boldsymbol{\beta}|_1$. Similar to the *least absolute shrinkage and selection operator* (LASSO) by Tibshirani (1996), this approach allows to obtain a sparse model through the l_1 norm dependent on the penalty term λ_1. The quadratic part of the penalty, on the other hand, encourages a grouped selection of correlated variables. For a given λ_2, we then obtain the first 30 of the $\breve{M} = 180$ variables, for which the elastic net method provides non-zero parameter estimates.[16] From this subset of 30 predictors, we then determine the factors and estimate the (augmented) dynamic factor model as in (3.3).

For these three set-ups (ARX, DFM, and DFM with elastic net preselection), we conduct a pseudo-out-of-sample forecast experiments based on data from 1975M1 to 2014M12 ($T = 480$) for forecast horizons of $h = 0, 1, 3, 6, 9, 12, 18, 24$ months. The $h = 0$ horizon gives the nowcast for the unpublished IPT realization. The first direct h-step forecasts are carried out in 1983M7 for all horizons from 1983M7 to 1985M7 ($\tau_0 = 103 - \max(l_0, l_i)$ observations).[17] At each iteration, the information set is extended by one month (recursive approach), all 248 models are re-estimated, the lag lengths P and Q_m are updated and all h-step forecasts are made.

3.3.2 Forecast evaluation

We evaluate the models' forecast performance by simple (relative) forecast accuracy and forecast rank measures, as well as by pairwise and joint tests on equal predictive accuracy. For this, we first obtain the forecast errors of each model based on (revised) data of 2014M12.

The first measure we construct is the root mean square prediction error (RMSPE) given by

$$RMSPE_{m,h} = \sqrt{\frac{\sum_{t=\tau_0+h}^{T} e_{h,m,t}^2}{T - \tau_0 - h + 1}}, \tag{3.5}$$

[16] We implement the elastic net estimation using the *glmnet* and *elasticnet* packages provided in *R* and determine the λ_2 penalty term using a 1,000-fold cross-validation approach. The choice of 30 series follows Bai and Ng (2008) but we find our results to be robust when increasing this choice to 50.

[17] As in the related literature, we only perform direct forecasts for $h > 0$. Also, predicting interim values up to h is not feasible for the bubble indicators.

where $e_{h,m,t} = (\hat{y}^h_{m,t+h} - y_{t+h})$ is the h-step forecast error made by model m in forecast period t. In order to evaluate the forecast accuracy of each candidate predictor, we employ the RMSPE ratio defined as $RR_{m,h} = \frac{RMSPE_{m,h}}{RMSPE_{0,h}}$ where model $m = 0$ is the AR (or DFM) forecast. A $RR_{m,h} < 1$ thus implies a better predictive accuracy of the ARX model compared to the benchmark, non-augmented AR model. Further, we provide the RMSPE rank of each model m at each forecast horizon h. The best model is thus the model that achieves the overall smallest RMSPE.

To test whether the gains or losses from augmenting the simple AR or DFM models by the alternative predictors are statistically significant, we apply the modified test statistic of Clark and West (2007). As the conventional test by Diebold and Mariano (1995) (DM test), this pairwise comparison assesses the null hypothesis that two models perform equally well. For this, we define the time t loss difference between a candidate model m for horizon h and the benchmark as $d_{0,m,h,t} = e^2_{h,0,t} - e^2_{h,m,t}$. The null hypothesis of equal predictive ability is thus $H_0 : E(d_{0,m,h,t}) = 0$. The standard DM test can, however, not be applied here since the benchmark model is nested in the augmented model m, such that, under the null hypothesis, the additional predictor variables do not provide new information to the benchmark forecast. Yet, their estimation introduces additional noise to the forecast of the alternative model, such that it must provide less accurate forecasts than the benchmark if the null hypothesis is true. Consequently, the DM test of equal predictive accuracy must be adjusted to account for the fact that the two models are nested. Clark and West (2007) show that this adjustment is given by

$$CW_{m,h} = \frac{\bar{d}_{h,m} - \bar{a}_{h,m}}{\hat{V}(\bar{d}_{h,m} - \bar{a}_{h,m})} \tag{3.6}$$

where $\bar{d}_{0,m,h}$ is the estimated mean of $d_{0,m,h,t}$ and $\hat{V}(.)$ is the estimated long-run variance of the adjusted $d_{0,m,h,t}$, with $\bar{a}_{h,m} = \frac{1}{T-\tau-h+1} \sum_{t=\tau}^{T-h} (\hat{y}^h_{0,t+h} - \hat{y}^h_{m,t+h})^2$.

While the Clark and West (2007) test allows to compare the accuracy of a candidate model to a predetermined benchmark, it does not identify the overall best models as it is sensitive to the choice of the benchmark. Therefore, we also identify the 25 percent Model Confidence Set (MCS) of Hansen et al. (2011).[18] This

[18]The assumptions for the MCS require forecasts to be obtained from a rolling window estimation approach. Thus, we present results for a pseudo-MCS here. For the forecast experiment of

approach aims to identify a subset of *superior* models $\mathcal{M}^\star$ from the set of all models M. Here, the set of superior models outperform all inferior models in terms of forecasting accuracy. This subset is hence defined as

$$\mathcal{M}^\star \equiv \{m \in M : E(d_{m,j,h,t}) \leq 0 \; \forall \; j \in M\} \tag{3.7}$$

where $E(d_{m,j,h,t})$ is the pairwise loss differential between models m and j based on squared forecast errors as above. Thus, this procedure carries out one-by-one comparisons of all models, and attempts to successively eliminate the worst performing model until the null hypothesis of equal predictive performance among all remaining models cannot be rejected at confidence level α. In the best case, when the data is very informative, the MCS consists of one unique model. When the data is uninformative, the procedure yields a subset of many, or even all, models.

Specifically, the MCS is obtained through the following steps with the candidate MCS $\mathcal{M} = M$ at the first iteration.

1. The null hypothesis of equal predictive accuracy, i.e. $H_{0,\mathcal{M}} : E(d_{m,j,h,t}) \leq 0 \; \forall \; m, j \in \mathcal{M}$ is assessed at significance level α.

2. If H_0 is rejected, the worst-performing model is eliminated from $\mathcal{M}$.

3. This is repeated until the null hypothesis cannot be rejected anymore. The set of remaining models is then defined as the MCS $\mathcal{M}^\star_{100(1-\alpha)}$, where we set $\alpha = 0.25$.

We test for equal predictive accuracy in step one using the $T_{max,\mathcal{M}}$ statistic. For this, we define the t-statistic

$$t_{m,.,h} = \frac{\bar{d}_{m,.,h}}{\sqrt{\hat{V}(\bar{d}_{m,.,h})}}, \tag{3.8}$$

where $\bar{d}_{m,.,h}$ is the horizon h loss of model m relative to the average across models in $\mathcal{M}$, i.e. $\bar{d}_{m,.,h} = n^{-1} \sum_{j \in \mathcal{M}} \bar{d}_{m,j,h}$, with $n = |\mathcal{M}|$. This t-statistic can then be

Stock and Watson (1999), Hansen et al. (2011) show that the results for the pseudo-MCS under the recursive approach are very similar to the MCS obtained from a rolling window approach. For our study, however, the predictive ability of the binary bubble indicators is limited when conducting a rolling window estimation due to the sparsity of the bubble signals.

used to assess the null $H_{m,.,h} : E(\bar{d}_{m,.,h}) = 0$. Hansen et al. (2011) show that the null hypothesis $H_{0,\mathcal{M}}$ defined in step one is equivalent to $\{H_{m,.,h} \; \forall \; m \in \mathcal{M}\}$, which extends to $\{E(\bar{d}_{m,.,h}) \leq 0 \; \forall \; m \in \mathcal{M}\}$. Thus, the null hypothesis $H_{0,\mathcal{M}}$ can be tested using the statistic

$$T_{max,\mathcal{M}} = \max_{m \in \mathcal{M}} t_{m,.,h}, \tag{3.9}$$

where the asymptotic distributions of these test statistics can be obtained using bootstrap methods.[19] If the null $H_{0,\mathcal{M}}$ is rejected, the worst-performing model must be eliminated. We identify this model by the elimination rule $e_{max,\mathcal{M}} = \arg\max_{m \in \mathcal{M}} t_{m,.,h}$. This model has the largest standardized excess loss relative to the average of all other models and thereby contributes the most to the test statistic.

Finally, we will not only assess the models' overall performance, but also their forecast ability for output growth during expansion and recession periods. For this subsample analysis, we evaluate the relative RMSPE ratios and the models' RMSPE ranks based on forecast errors occurring during expansion and recession periods as defined by the NBER.[20] Since this yields series of nonconsecutive forecast errors, the Clark and West (2007) test and the MCS evaluation are not defined. Figure 3.A.5 depicts the recession periods and the IPT series.

3.4 Results

3.4.1 Overall predictive accuracy

Table 3.4.1 displays the main results from our real-time pseudo-out-of-sample forecast experiment for all forecast horizons $h = 0, 1, 3, 6, 9, 12, 18, 24$. The second row displays the number of models included in the MCS $\mathcal{M}^*_{75}$. For most horizons except for the nowcast (h=0), the MCS eliminates at least 60% of the models. The fewest number of survival models are obtained for the one-year forecast horizons, for which 31 models are found to be superior to all excluded models.

[19] We implement this procedure using the MFE Toolbox implemented in Matlab by Kevin Sheppard, available at `https://www.kevinsheppard.com/MFE_Toolbox`.

[20] These dates are available at `www.nber.org/cycles.html`. The results for the subsample performance also hold when expansion and recession periods are defined based on the turning-point algorithm by Bry and Boschan (1971) and Harding and Pagan (2002) for the target variable IPT.

Table 3.4.1: ARX forecasts: relative RMSPE, forecast rank, and MCS

Horizon	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**	**Mean**
MCS	154	61	81	95	36	31	63	106	80
AR	7.368 (98)	5.602 (107)	4.738 (99)	4.513 (115)	4.355 (130)	4.166 (131)	3.799 (127)	3.417 (138)	4.745 (124)
DFM	.983	.981	1.008	1.032	1.021	1.030	1.012	1.034	1.009
DFM ELA	.952	.918	.936	1.003	.983	.979	1.004	.976	0.965
				Best non-bubble indicator					
NAPM	.932 (1)								.984 (97)
NAPM		.887 (1)							.984 (97)
DY.D1			.916 (1)						.945 (5)
DY.D1				.894 (1)					.945 (5)
ENRG.P12					.899 (1)				.933 (1)
ENRG.P12						.905 (1)			.933 (1)
ENRG.P12							.892 (1)		.933 (1)
T1YFFM								.883 (1)	.951 (12)
				Stock price bubbles					
HP_{rec}	.995 (45)								1.217 (239)
HP_{rol}	.995 (48)								1.207 (235)
$PWY11^{i}_{rec}$	.999 (80)		.982 (36)	.973 (38)	.943 (30)		.929 (52)	.961 (84)	.966 (42)
$PWY11^{r}_{rec}$	1.001 (107)		.989 (53)	.987 (83)					.977 (71)
$PWY11^{i}_{rol}$	**.990 (21)**	**.987 (34)**	**.967 (11)**	*.956 (13)*	*.927 (11)*	*.925 (10)*	*.909 (22)*	.922 (38)	.948 (8)
$PWY11^{r}_{rol}$	.993 (32)	.997 (83)	.980 (31)	*.966 (18)*	*.931 (12)*	*.927 (12)*	.909 (26)	.929 (52)	.954 (16)
$PSY15^{i}$	.997 (59)		*.975 (20)*	**.956 (12)**	**.922 (6)**	**.915 (2)**	**.895 (2)**	**.916 (29)**	**.947 (7)**
$PSY15^{r}$	1.008 (170)		.989 (51)	.974 (41)	*.938 (19)*		.920 (40)	.926 (46)	.962 (34)
$B^{Comb}_{\tau,1}$	1.006 (157)								1.234 (241)
$B^{Comb}_{\tau,2}$	1.002 (113)							.986 (116)	1.213 (237)
$B^{Comb}_{\tau,3}$	*.989 (19)*	**.981 (22)**	**.957 (5)**	**.947 (8)**	**.920 (4)**	**.917 (3)**	**.905 (10)**	.925 (45)	**.943 (4)**
$B^{Comb}_{\tau,4}$	**.988 (14)**	*.983 (25)*	*.960 (6)*	*.952 (9)*	*.922 (7)*	*.920 (6)*	*.908 (21)*	.927 (47)	.945 (6)
$B^{Comb}_{\tau,5}$	.993 (33)	.991 (43)	*.963 (7)*	*.955 (11)*	*.926 (10)*	*.923 (8)*	*.910 (27)*	.927 (48)	.948 (9)
$B^{Comb}_{\tau,6}$	.997 (57)	.988 (35)	*.967 (10)*	*.954 (10)*	*.923 (8)*	*.927 (13)*	.913 (31)	.924 (42)	.949 (11)
$B^{Comb}_{\tau,7}$	.998 (75)	.990 (41)	*.973 (18)*	*.968 (22)*	*.938 (20)*		.916 (36)	**.923 (41)**	.955 (20)
$B^{Comb}_{\tau,8}$	.999 (82)	.991 (42)	*.973 (17)*	*.967 (20)*	*.937 (18)*		.917 (38)	.925 (43)	.956 (22)
B^{RW}_{τ}	.994 (41)	.994 (54)	*.974 (19)*	*.966 (17)*	*.936 (16)*	*.934 (19)*	.915 (35)	.939 (56)	.957 (25)
				House price bubbles					
HP_{rec}									1.010 (146)
HP_{rol}	.999 (79)		**.983 (37)**	**.974 (39)**	**.935 (14)**	**.930 (17)**	**.902 (5)**	*.907* (22)*	**.953 (13)**
$PWY11^{i}_{rec}$	**.996 (49)**		.990 (56)	.979 (57)			*.905 (11)*	**.900* (13)**	.956 (21)
$PWY11^{r}_{rec}$	.998 (73)						.925 (47)	.916* (28)	.966 (43)
$PWY11^{i}_{rol}$	1.004 (143)								.979 (78)
$PWY11^{r}_{rol}$	1.007 (167)						.920 (41)	.918* (33)	.976 (63)
$PSY15^{i}$	.998 (74)		.995 (75)				.913 (32)	*.908* (24)*	.963 (36)
$PSY15^{r}$	.997 (65)							.918* (35)	.970 (51)
$B^{Comb}_{\tau,1}$	**.993 (39)**	**.991 (45)**	.990 (57)	.982 (64)			.931 (57)	.918 (34)	.966 (41)
$B^{Comb}_{\tau,2}$	.999 (85)		.994 (65)	.984 (71)				.916* (30)	.970 (48)
$B^{Comb}_{\tau,3}$	1.000 (94)						.922 (45)	**.913* (27)**	.967 (45)
$B^{Comb}_{\tau,4}$	1.004 (135)								.980 (81)
$B^{Comb}_{\tau,5}$			.995 (71)	.985 (76)					.977 (66)
B^{RW}_{τ}	.997 (66)		**.979 (30)**	**.978 (55)**	**.938 (21)**	**.935 (25)**	**.915 (34)**	.923* (40)	**.958 (29)**

Notes: The table shows the forecast performance of the ARX models. The second line displays the number of models included in the MCS. The third line shows the RMSPE of the AR model and its horizon-specific rank among all models in parentheses. The fourth and fifth line display the relative RMSPE of the DFM and the DFM with elastic-net preselection. The top panel shows the relative RMSPE of the best performing non-bubble indicator for each forecast horizon. The middle panel shows the same results for all stock price bubble indicators; the bottom panel gives the results for the house price bubble indicators. Bubble indicators excluded from the MCS are not displayed for ease of presentation. For forecasts denoted with '*', the hypothesis of equal predictive performance of the AR and the candidate predictor is rejected at the ten percent significance level by the Clark and West (2007) test. The best indicators at each horizon for each class (stock vs. house, individual vs. combination) are highlighted in bold. Values in italics are among the best ten percent of all models. The last column shows the mean value (relative) RMSPE across all evaluated horizons.

The third to fifth rows display the absolute RMSPE of the AR model, as well as the relative RMSPE ratios for the DFM and the DFM with elastic net preselection (DFM ELA). Also, the forecast rank of the AR is given in parentheses. Overall, we find that the AR model performs around the median of all models with forecast ranks of 98 to 138 and is excluded from the MCS for all horizons except for the nowcast $h = 0$. This implies that on average across all horizons half of the ARX models outperform the simple AR forecast. Furthermore, while the DFM based on all stationary transformations of all predictors does not outperform the AR model except for $h = 0$ and $h = 1$, the DFM with elastic net preselection strongly outperforms the AR except for two horizons ($h = 6$ and $h = 18$).

Turning to the individual indicators, the top panel displays the relative RMSPE of the best non-bubble predictor for each forecast horizon. For short horizons, the survey-based purchasing managers index (ISM Manufacturing: PMI Composite Index, NAPM) performs best, outperforming the AR model by 6.8% to 11.3%. Across all horizons, however, this index is only the 97th best model (see last column). In contrast, the non-bubble indicators performing best at medium- to long-term horizons are also among the best twelve predictors overall. These are the dividend yield (in first differences, DY.D1), the year-on-year percentage change in the World Bank Energy Index (ENRG.P12), and the term spread between one-year Treasuries and the Federal Funds Rate (TY1FFM).[21]

While there is always one non-bubble indicator that outperforms all bubble indicators, several bubble indicators are competitive in terms of forecast accuracy. The middle panel shows the performance of the stock price bubble indicators, where we suppress entries that are excluded from the MCS. Overall, we find that at least eight out of 17 indicators are included in the MCS. Among these, the $PWY11^i_{rol}$ and $PSY15^i$ indicators, as well as the combinations $\mathcal{B}^{Comb}_{\tau,\kappa}$ with $\kappa = 3, \ldots, 6$ are the best performing indicators, ranking fourth to eleventh overall. Among the individual indicators, the $PWY11^i_{rol}$ provides the best forecasts for the short-term horizons up to $h = 3$, while the $PSY15^i$ is among the overall best indicators for $h = 6$ and

[21] Among the other overall best indicators are the dividend yield in levels (DY, 2nd overall), housing starts (HSTARTS.P12, 3rd), and total capacity utilization (CUT.D2, 10th). Further, we find that only few non-bubble ARX models are excluded from the MCS for all horizons. These are denoted by 'x' in Table 3.B.2. The share of excluded models is similar for all transformations, such that several trending level variables are included in the MCS for at least some horizons. This indicates that pretesting for and excluding $I(1)$ predictors may discard relevant predictors.

longer. Nonetheless, both are outperformed for horizons up to $h = 12$ by the simple combinations $\mathcal{B}^{Comb}_{\tau,\kappa}$ with $\kappa = 3, 4$. It is noteworthy that these best individual and combination indicators also provide the most plausible bubble signals as demonstrated in Figures 3.A.1 and 3.A.2. In contrast, the HP indicators rank among the worst indicators overall. Nonetheless, they do provide some valuable information when combined with the unit root indicators to $\mathcal{B}^{Comb}_{\tau,\kappa}$ with $\kappa \geq 3$.

The bottom panel shows the results for the house price bubble indicators. Generally, these indicators are frequently excluded from the MCS. Across all horizons, however, these signals help to outperform the AR forecast. For the long-term forecast of $h = 24$, they rank among the best 25 models, and their predictive gains against the AR model are statistically significant at the ten percent level. At this horizon, they also outperform the stock price bubble indicators.

Since the DFM with elastic net preselection outperforms the AR models and all bubble indicator-augmented models at most horizons, we assess the performance of the bubble-augmented DFM in Table 3.B.3. Again, the best stock price bubble indicators provide more accurate forecasts than the DFM with elastic net preselection. We thus conclude that these indicators provide additional predictive value for output growth even when all information from a large set of conventional macroeconomic and financial predictors is already considered. This also suggests that the bubble episodes detected by the best indicators are systematically related to higher or lower future output growth.

3.4.2 Predictive accuracy in expansion and recession periods

As discussed in the introduction, the additional predictive value of such bubble indicators may result from the causal effects that pronounced asset price appreciations have on the economy by intensifying expansions during their boom and by causing recessions after their burst. We therefore explore from which periods the predictive gains of the bubble indicators stem.

For this, we replicate the previous results but condition on NBER expansion and recession periods. Table 3.B.4 displays the forecast performance of all models for NBER expansion periods. During expansions, the AR forecast is more accurate in absolute and in relative terms. Specifically, the AR forecast now ranks 71st overall (see the last column) versus 124th in the full-sample analysis. For the three-

month horizon, only 16 models outperform the AR forecast. Again we display the best performing non-bubble indicators in the upper panel. Since most periods are expansion periods, the best performing non-bubble predictors are almost identical to the ones for the full-sample evaluation. However, their overall forecast ranks and their relative gains to the AR deteriorate. Moreover, we note that the DFM and the DFM ELA only help to improve the immediate forecasts.

In contrast, the best performing bubble indicators from the full-sample analysis again help to improve the forecast accuracy of the AR model. Specifically, stock price bubble indicators now provide the overall seven best forecasts (see last column). Again, these are the $PWY11^{i}_{rol}$ and $PSY15^{i}$ indicators, as well as the combinations $\mathcal{B}^{Comb}_{\tau,\kappa}$ with $\kappa = 3, \ldots, 5$. However, also the $PWY11^{r}_{rol}$ and the econometric combination $\mathcal{B}^{RW}_{\tau}$ are among the best seven predictors. Thus, we conclude that the predictive gains of these binary indicators stems mostly from an upward correction of the AR forecast during expansion periods.[22] Since the DFM hardly outperforms the AR benchmark, we do not show the results for the augmented DFM but only note that most bubble indicators again help to improve forecasts from the DFM and with DFM with elastic net preselection.

Table 3.B.5 shows the forecast performance during recession periods. For these periods, the AR model performs considerably worse, both in absolute and in relative terms. On average, 190 indicators help to improve the forecasts of the AR model. The results for the best-non bubble indicators emphasize that their predictive value is considerably higher during recession periods. In contrast, while the bubble indicators continue to frequently outperform the AR model and rank among the best 100 models (top 40%), their predictive value is smaller than the one of the best macroeconomic and financial predictors. However, it is important to note that this finding is obtained from 34 months of recessionary periods only. Moreover, the best bubble indicators from the expansion or full-sample analysis again outperform the AR forecast.

Furthermore, predictive gains during recessions can also be obtained from condensing the information from the best macroeconomic and financial predictors by

[22] We find that the full-sample parameter estimates $\gamma_{m,h,q}$ for horizons $h \leq 9$ are positive and significant for all best performing bubble indicators (see Table 3.B.7). For $h > 9$, the signs of the most plausible indicators ($PWY11^{i}_{rol}$ and $\mathcal{B}^{Comb}_{\tau,3}$, $\mathcal{B}^{Comb}_{\tau,4}$) revert, presumably capturing the ensuing recession.

means of the DFM and the DFM with elastic net preselection, in particular. This latter model now outperforms the AR model for all horizons. As shown in Table 3.B.6, this model can, however, be further improved by augmenting it by the $\mathcal{B}_{\tau,3}^{Comb}$ combination indicator applied to stock prices or the $PSY15^r$ indicator applied to house prices. All in all, we thus conclude that real-time bubble indicators as proposed by Phillips et al. (2011, 2015) and the combination indicators developed in Chapter 2 provide a promising tool to detect asset (and in particular stock) price bubbles in real-time, and that these signals can further be employed to help predict future real economic activity.

3.5 Conclusion

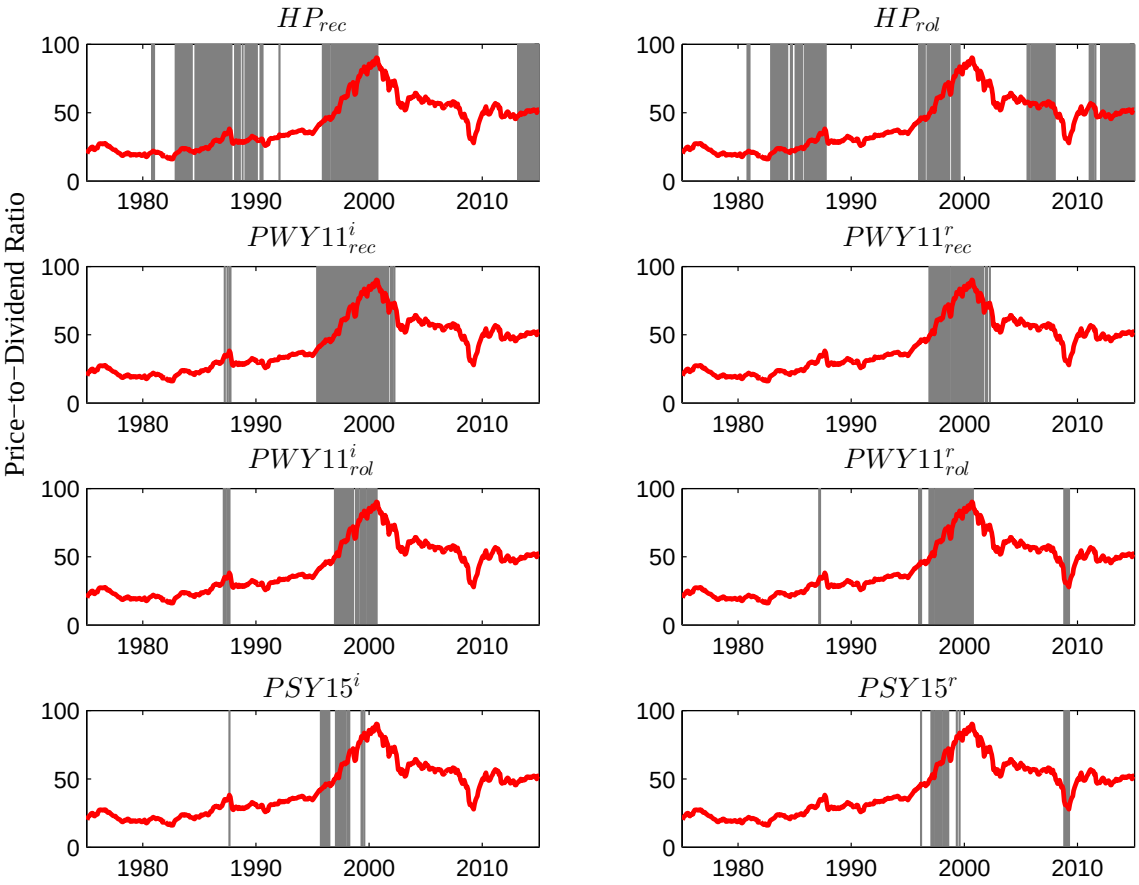
In this paper, we explore whether available monitoring indicators for asset price bubbles can help to improve forecasts for output growth in real-time. For this, we compare them to a comprehensive set of 216 models based on macroeconomic and financial indicators that are commonly used to predict real economic activity or inflation. In particular, our set of real-time bubble indicators makes use of monitoring tests for explosive growth in stock or house prices, of deviations from a HP-filtered trend, or of a simple counting or econometric combinations of these tests.

We first show that several indicators provide plausible signals for stock price bubbles in real-time. Subsequently, we then show that these same bubble indicators also help to improve output forecasts from an autoregressive model, and are among the ten best overall predictors when compared to the large set of rival predictors. Finally, we show that these predictive gains stem mostly from expansion periods in output growth, but forecasts can also be improved during recession periods when combining a dynamic factor model of the best conventional predictors with the most promising bubble indicators. We therefore argue that real-time indicators for asset price bubbles can provide useful supplementary information for output forecast.

Appendix

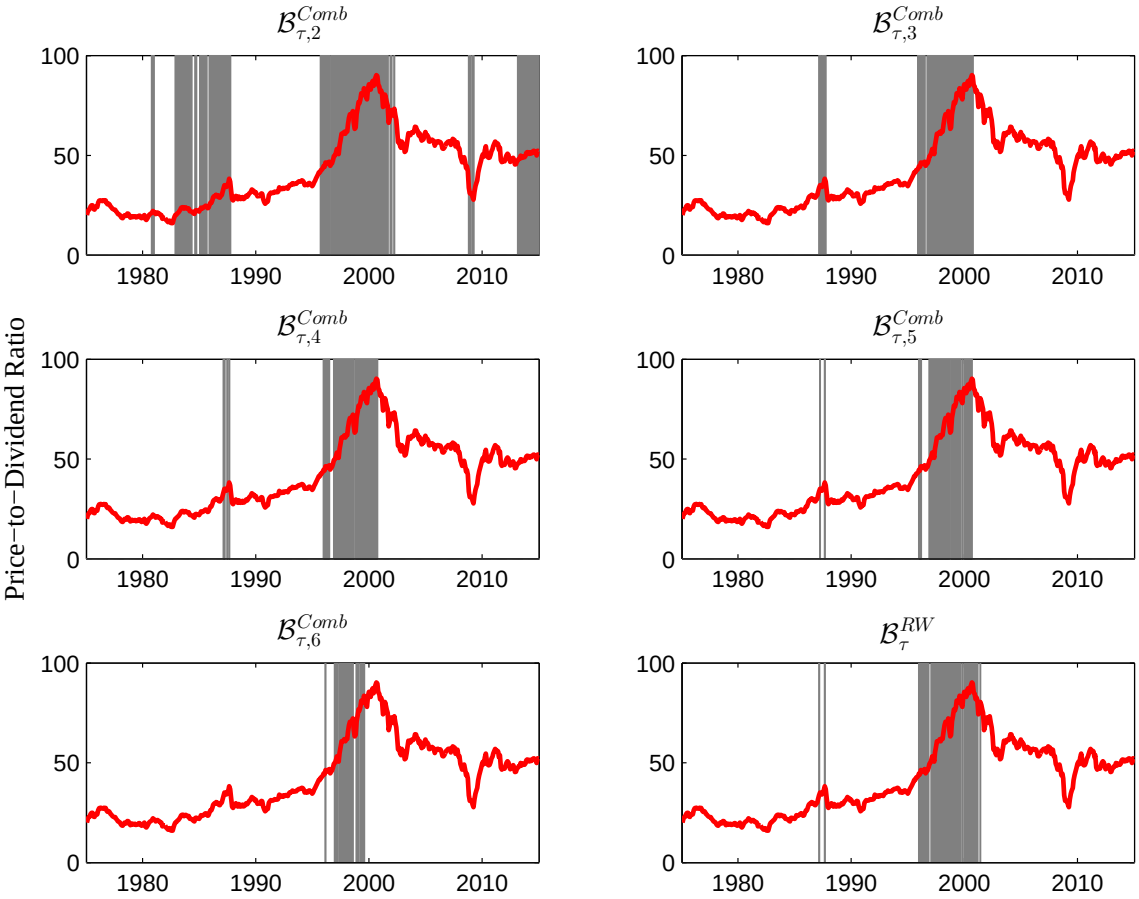
3.A Figures

Figure 3.A.1: Bubble periods in the S&P 500 as detected by individual indicators



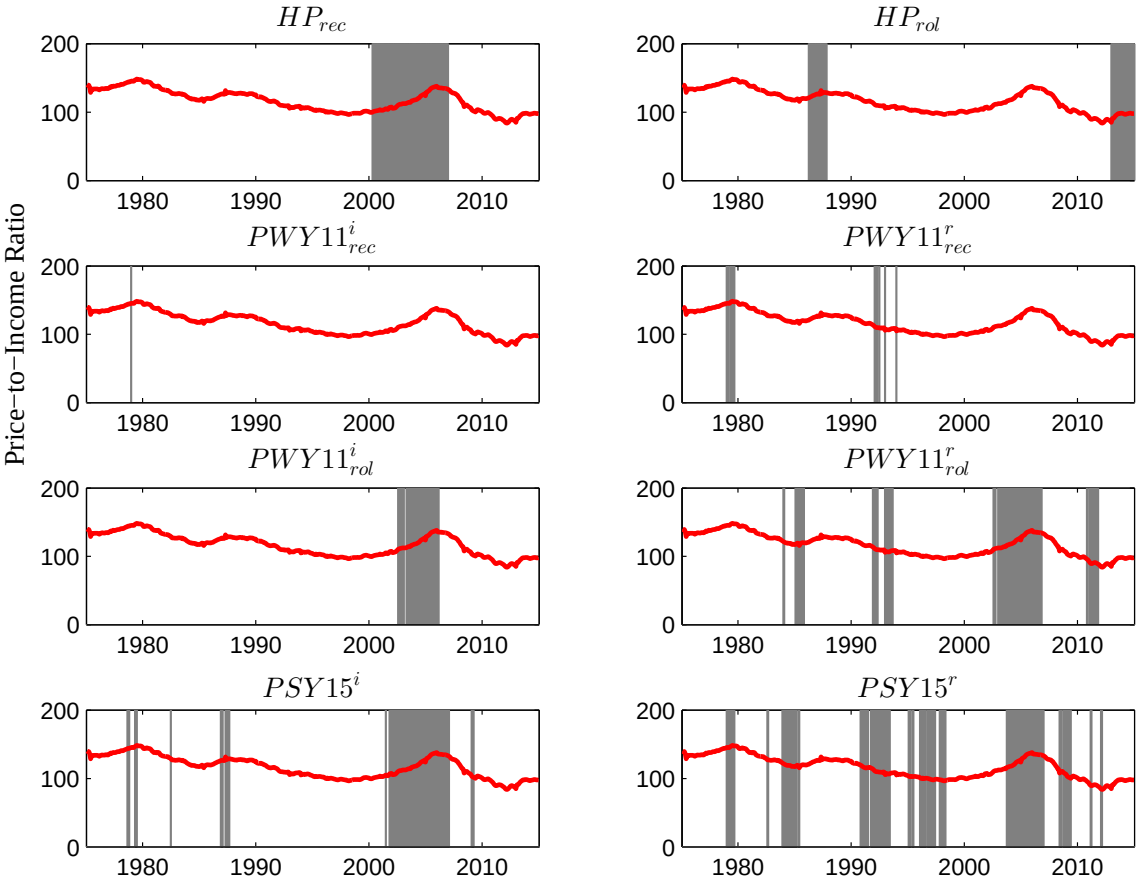
Notes: The figure shows the bubble periods in the S&P 500 as detected by the real-time individual indicators. The top panel shows the indicators based on the HP-filter (left: recursive, right: rolling). The bottom three panels show the signals from the Phillips et al. (2011, 2015) indicators. Here, the left panels show the indicators when applied to the price and dividend series individually. The right bottom panels show the indicators when applied to the price-to-dividend (PtD) ratio. The solid line plots the PtD ratio.

Figure 3.A.2: Bubble periods in the S&P 500 as detected by combination indicators



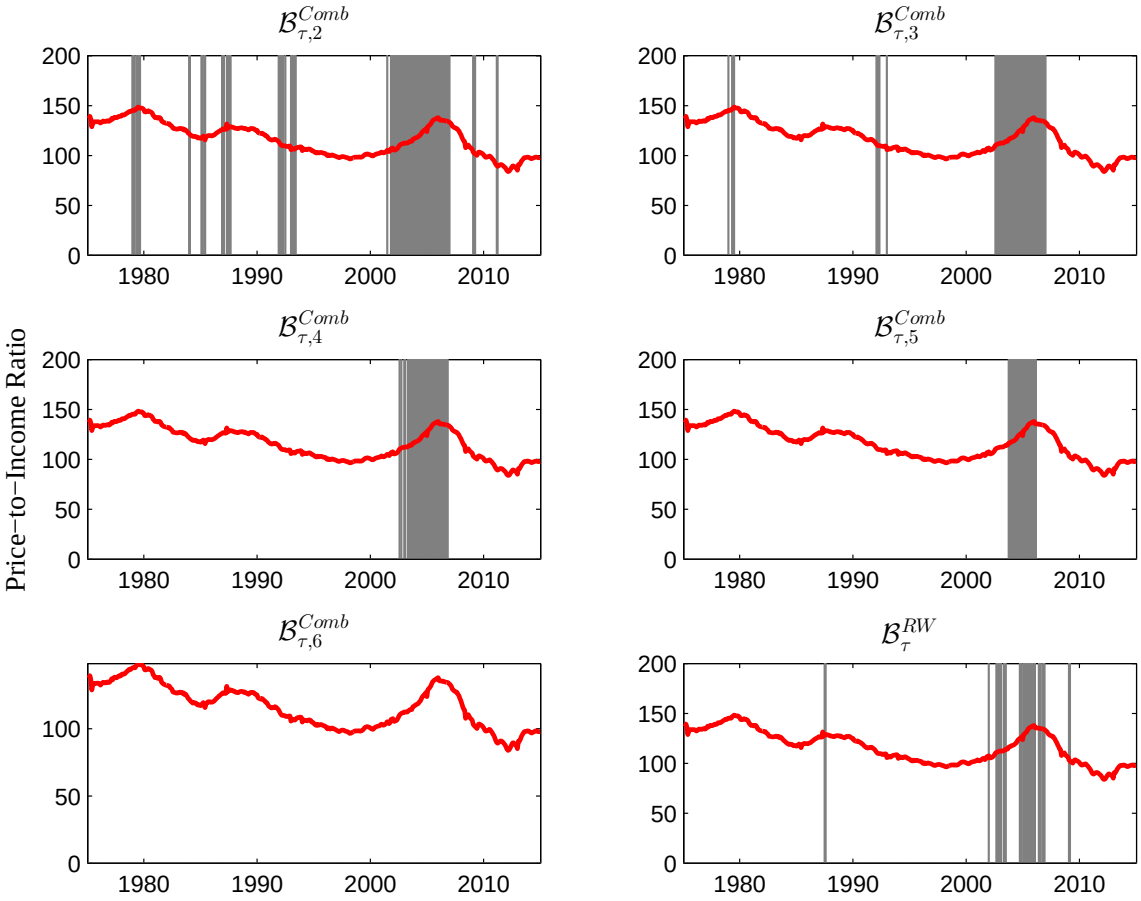
Notes: The figure shows the bubble periods in the S&P 500 as detected by the counting combination indicators $\mathcal{B}^{Comb}_{\tau,\kappa}$ for different threshold levels κ, and by the econometric combination $\mathcal{B}^{RW}_{\tau}$ (bottom right panel). See Figure 3.A.1 for further notes.

Figure 3.A.3: Bubble periods in house prices as detected by individual indicators



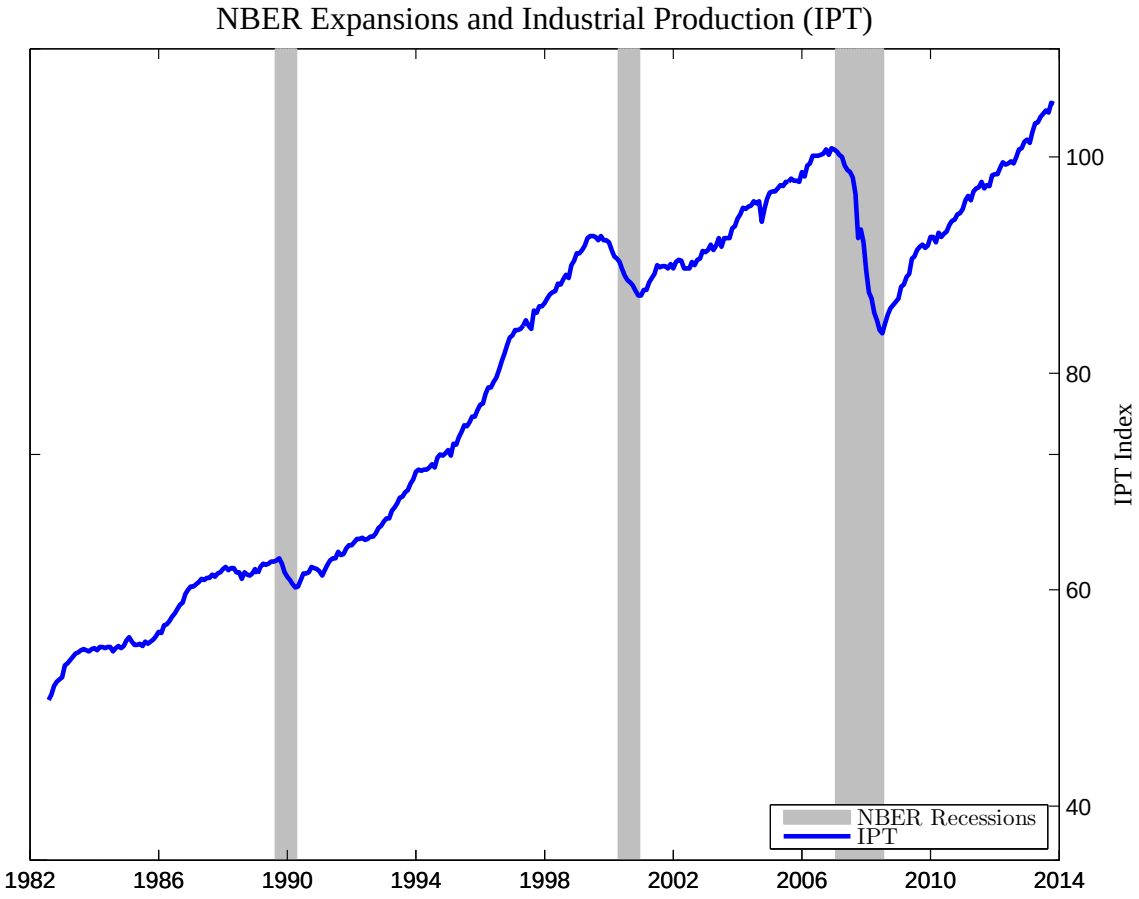
Notes: The figure shows the bubble periods in U.S. house prices as detected by the individual indicators. The indicators are ordered as in Figure 3.A.1. The left bottom three panels show the Phillips et al. (2011, 2015) indicators applied to the real price and disposable income series individually. The right bottom panels shows the indicators applied to the price-to-income (PtI) ratio. The solid line plots the PtI ratio.

Figure 3.A.4: Bubble periods in house prices as detected by combination indicators



Notes: The figure shows the bubbles periods in U.S. house prices as detected by the counting combination indicators $\mathcal{B}^{Comb}_{\tau,\kappa}$ for different threshold levels κ, and by the econometric combination $\mathcal{B}^{RW}_{\tau}$ (bottom right panel). See Figure 3.A.3 for further notes.

Figure 3.A.5: NBER recession dates and total industrial production



Notes: The figure shows U.S. recession periods (grey bars) as defined by the NBER and the evolution of the total industrial production index (IPT, right axis).

3.B Tables

Table 3.B.2: Data: definitions, sources, lags and revisions, and transformations

Variable	Identifier	Source	Lag	SA	REV	L	D1	D2	P1	P12
Real economic variables										
Industrial Production Index: Total	IPT	RTDS	1	yes	yes	y				
Industrial Production Index: Manufacturing	IPM	RTDS	1	yes	yes	y	x			
ISM Manufacturing: PMI Composite Index	NAPM	FRED	1	-	-			x		
Capacity Utilization Rate: Manufacturing	CUM	RTDS	1	yes	yes	x, y	x		x	
Capacity Utilization Rate: Total	CUT	RTDS	1	yes	yes	y				
Nonfarm Payroll Employment	EMPLOY	RTDS	1	yes	yes	y				
Indexes of Aggregate Weekly Hours: Total	H	RTDS	1	yes	yes	x, y				
Indexes of Aggregate Weekly Hours: Goods-Producing	HG	RTDS	1	yes	yes	x, y				
Indexes of Aggregate Weekly Hours: Service-Producing	HS	RTDS	1	yes	yes	x, y				
Housing Starts	HSTARTS	RTDS	1	yes	yes	y				
Prices and Energy										
Consumer Price Index all urban consumers	CPI	FRED	1	-	-	y				
Core Consumer Price Index	PCPIX	FRED	1	-	-	x, y		x	x	x
Producer Price Index (Finished Goods)	PPI	BLS	1	-	-	y	x		x	
Core Producer Price Index (Finished Goods)	PPPIX	BLS	1	-	-	x, y	x		x	
Nominal House Price Index	HPI	Shiller	2	-	-	y		x		
Spot Oil Price: West Texas Intermediate	OILPRICE	FRED	0	-	-	y	x	x		
Crude Oil, Average of U.K. Brent, Dubai & West Texas Intermediate, End of Period, USD	OILAVG	MB	0	-	-	y	x		x	
Commodity Indices, World Bank, Non-Energy Index, End of Period, USD	NONENRG	MB	0	-	-	y	x			
Commodity Indices, World Bank, Metals & Minerals Index, End of Period, USD	METAL	MB	0	-	-	y				
Commodity Indices, World Bank, Food Index, End of Period, USD	FOOD	MB	0	-	-	y				
Commodity Indices, World Bank, Energy Index, End of Period, USD	ENRG	MB	0	-	-	y				

Data: definitions, sources, lags and revisions, and transformations (ctd.)

Variable	Identifier	Source	Lag	SA	REV	L	D1	D2	P1	P12
Financial variables										
Nominal S&P 500 Index	SP500	Shiller	0	-	-	y	x			
S&P 500 Dividends	DIV	Shiller	1	-	-	y	x			
S&P 500 Earnings	EAR	Shiller	4	-	-	x		x	x	x
S&P 500, Index, Dividend Yield, Average of Period	DY	MB	0	-	-	y		x		
S&P 500, Index, P/E Ratio, Average of Period	PE	MB	0	-	-	x	x	x	x	x
Effective Federal Funds Rate	FEDFUNDS	FRED	0	-	-	y				x
3-Month Treasury Bill: Secondary Market Rate	TB3MS	FRED	0	-	-	y			x	x
3-Month Treasury Bill Minus Federal Funds Rate	TB3SMFFM	FRED	0	-	-				a	a
6-Month Treasury Bill: Secondary Market Rate	TB6MS	FRED	0	-	-	y				
6-Month Treasury Bill Minus Federal Funds Rate	TB6SMFFM	FRED	0	-	-				a	a
1-Year Treasury Constant Maturity Rate	GS1	FRED	0	-	-	y				x
1-Year Treasury Constant Maturity Minus Federal Funds Rate	T1YFFM	FRED	0	-	-				a	a
5-Year Treasury Constant Maturity Rate	GS5	FRED	0	-	-	y			x	x
5-Year Treasury Constant Maturity Minus Federal Funds Rate	T5YFFM	FRED	0	-	-				a	a
10-Year Government Bond Interest Rate	GS10Y	Shiller	0	-	-	y		x		x
10-Year Treasury Constant Maturity Minus Federal Funds Rate	T10YFFM	FRED	0	-	-				a	a
Moody's Seasoned Aaa Corporate Bond Yield	AAA	FRED	0	-	-	y				x
Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate	AAAFFM	FRED	0	-	-	x			a	a
Moody's Seasoned Baa Corporate Bond Yield	BAA	FRED	0	-	-	x, y	x	x	x	x
Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate	BAAFFM	FRED	0	-	-	x			a	a
3-Month AA Nonfinancial Commercial Paper Rate	CPNF3M	MB	0	-	-	y				x

Data: definitions, sources, lags and revisions, and transformations (ctd.)

Variable	Identifier	Source	Lag	SA	REV	L	D1	D2	P1	P12
Exchange rates										
Trade Weighted U.S. Dollar Index: Major Currencies	TWEXMMTH	FRED	0	-	-	x, y				x
Canada / U.S. Foreign Exchange Rate	EXCAUS	FRED	0	-	-	y				
Japan / U.S. Foreign Exchange Rate	EXJPUS	FRED	0	-	-	x, y				
Switzerland / U.S. Foreign Exchange Rate	EXSZUS	FRED	0	-	-	y				
U.S. / U.K. Foreign Exchange Rate	EXUSUK	FRED	0	-	-	y				

Notes: The table lists the set of macroeconomic and financial predictors used to forecast industrial production (IPT). The sources are the Real-Time Dataset for Macroeconomists from the Federal Reserve Bank of Philadelphia (RTDS), the Federal Reserve Economic Data from the Federal Reserve Bank of St. Louis (FRED), the Bureau of Labor Statistics (BLS), Macrobond (MB), and the stock and housing market data provided by Robert Shiller (Shiller) available at `http://www.econ.yale.edu/~shiller/data.htm`. Column 'Lag' displays the publication lag in months. 'SA' denotes which series are seasonally adjusted and 'REV' denotes which series are revised after initial publication. The last five columns indicate the different transformations of the raw data: level (L: x_t), first differences (D1: $\Delta x_t = x_t - x_{t-1}$), second differences (D2: $\Delta^2 x_t = \Delta x_t - \Delta x_{t-1}$), month-to-month percentage changes (P1: $1200\frac{x_t - x_{t-1}}{x_{t-1}}$) or year-on-year percentage changes (P12: $100\frac{x_t - x_{t-12}}{x_{t-12}}$). Entries denoted by 'a' are not included in the forecast exercise as the series are not defined at some margins. Entries denoted by 'x' are included in the forecast exercise but are found to be eliminated from the Model Confidence Set (MCS) for all forecast horizons. Entries denoted by 'y' are found to be non-stationary by at least two tests for stationarity and are thus excluded from the construction of factors for the DFM.

Table 3.B.3: Forecasts from augmented DFM with elastic-net preselection

Horizon	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**	**Mean**
DFM ELA	7.014 (13)	5.141 (23)	4.434 (17)	4.526 (8)	4.282 (7)	4.078 (19)	3.813 (14)	3.334 (10)	4.578 (14)
				Stock price bubbles					
HP_{rec}	.999 (10)	**.982 (1)**							1.063 (30)
HP_{rol}									1.063 (29)
$PWY11^{i}_{rec}$			.994 (9)						1.009 (21)
$PWY11^{r}_{rec}$									1.017 (26)
$PWY11^{i}_{rol}$		.993 (8)	**.989 (5)**	**.992 (5)**	.993 (5)	.986 (5)	.994 (5)		.994 (6)
$PWY11^{r}_{rol}$									1.009 (22)
$PSY15^{i}$				.994 (6)	**.989 (3)**	**.977 (2)**	**.982 (1)**	**.987 (1)**	**.992 (4)**
$PSY15^{r}$									1.015 (27)
$B^{Comb}_{\tau,1}$								.999 (9)	1.076 (32)
$B^{Comb}_{\tau,2}$									1.071 (31)
$B^{Comb}_{\tau,3}$		**.986 (2)**	**.976 (1)**	**.985 (1)**	**.985 (1)**	**.975 (1)**	**.989 (2)**		**.988 (1)**
$B^{Comb}_{\tau,4}$	.998 (8)	.987 (3)	.981 (2)	.989 (3)	.988 (2)	.980 (3)	.995 (6)		.991 (2)
$B^{Comb}_{\tau,5}$		.994 (10)	.982 (3)	.991 (4)	.990 (4)	.983 (4)	.995 (7)		.993 (3)
$B^{Comb}_{\tau,6}$	**.993 (2)**	.991 (4)	.991 (6)	1.000 (9)	.998 (6)	.989 (7)	.996 (8)	.995 (7)	.994 (5)
$B^{Comb}_{\tau,7}$	.997 (5)	.993 (7)	.994 (8)		1.001 (8)	.994 (10)		**.992 (2)**	.996 (7)
$B^{Comb}_{\tau,8}$	.998 (6)							.994 (4)	.998 (9)
B^{RW}_{τ}			.993 (7)	1.000 (10)	1.001 (9)				1.002 (16)
				House price bubbles					
HP_{rec}	.998 (7)								1.027 (28)
HP_{rol}							.993 (4)	**.994 (3)**	1.002 (19)
$PWY11^{i}_{rec}$					**1.001 (10)**		**.992 (3)**	.995 (5)	**.999 (12)**
$PWY11^{r}_{rec}$						**.993 (9)**		.998 (8)	1.002 (18)
$PWY11^{i}_{rol}$	**.994 (3)**	**.994 (9)**							1.001 (13)
$PWY11^{r}_{rol}$									1.015 (25)
$PSY15^{i}$			**.995 (10)**	**.999 (7)**					1.003 (17)
$PSY15^{r}$									1.013 (24)
$B^{Comb}_{\tau,1}$			**.988 (4)**	**.987 (2)**		**.987 (6)**			**.998 (8)**
$B^{Comb}_{\tau,2}$	.999 (9)								1.012 (23)
$B^{Comb}_{\tau,3}$						.991 (8)	**.998 (9)**	**.995 (6)**	1.000 (15)
$B^{Comb}_{\tau,4}$	**.993 (1)**	.993 (6)							1.000 (11)
$B^{Comb}_{\tau,5}$	.996 (4)	**.992 (5)**					.998 (10)		.998 (10)
B^{RW}_{τ}									1.005 (20)

Notes: The table shows the forecast performance of the (augmented) dynamic factor model (DFM) with elastic-net preselection for all forecast horizons given in the first line. The second line shows the RMSPE of the DFM with elastic-net preselection and its horizon-specific rank among all models in parentheses. The top panel shows the RMSPE (relative to the DFM ELA model) and rank of all stock price bubble indicators; the bottom panel shows the same results for the house price bubble indicators. Only the ten best bubble indicators at each horizon are shown for ease of presentation. See Table 3.4.1 for further notes.

Table 3.B.4: ARX forecasts: NBER expansion periods

Horizon	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**	**Mean**
Obs.	344	344	344	344	344	344	344	344	
AR	6.345 (100)	4.390 (69)	3.315 (17)	2.873 (32)	2.892 (97)	3.031 (96)	3.088 (117)	2.929 (123)	3.608 (71)
DFM	.993	1.014	1.063	1.173	1.148	1.098	1.083	1.078	1.067
DFM ELA	.977	.970	1.023	1.128	1.081	1.033	1.083	1.048	1.031
Best non-bubble indicator									
NAPM	.949 (1)								.994 (57)
NAPM		.902 (1)							.994 (57)
T1YFFM			.965 (1)						.968 (10)
DY.D1				.934 (1)					.973 (13)
DY.D1					.884 (1)				.973 (13)
DY.D1						.894 (1)			.973 (13)
ENRG.P12							.873 (1)		.967 (9)
T1YFFM								.882 (1)	.968 (10)
Stock price bubbles									
HP_{rec}	.992 (55)								1.339 (239)
HP_{rol}	.991 (50)								1.311 (233)
$PWY11^{i}_{rec}$	.986 (32)	1.004 (92)	1.026 (77)	1.026 (64)	*.929 (11)*	*.934 (11)*	.919 (34)	.968 (84)	*.977 (18)*
$PWY11^{r}_{rec}$	.993 (56)		1.046 (92)		.968 (59)	.968 (60)	.943 (68)	.983 (97)	.997 (64)
$PWY11^{i}_{rol}$	*.976 (8)*	.992 (36)	1.010 (50)	1.010 (50)	*.922 (9)*	*.917 (9)*	*.895 (7)*	.924 (29)	*.960 (4)*
$PWY11^{r}_{rol}$	**.975 (7)**	**.988 (32)**	**1.001 (22)**	1.012 (53)	*.926 (10)*	*.916 (8)*	**.891 (4)**	.925 (33)	**.958 (3)**
$PSY15^{i}$	.985 (27)		1.024 (72)	1.009 (46)	**.916 (6)**	**.914 (7)**	*.891 (5)*	**.920 (25)**	*.965 (7)*
$PSY15^{r}$	.991 (52)	.994 (42)	1.018 (62)	**1.008 (42)**	.951 (23)	.959 (43)	.920 (35)	.935 (49)	*.976 (16)*
$\mathcal{B}^{Comb}_{\tau,1}$	.997 (78)								1.350 (241)
$\mathcal{B}^{Comb}_{\tau,2}$	.989 (40)								1.316 (235)
$\mathcal{B}^{Comb}_{\tau,3}$	*.976 (9)*	**.985 (23)**	**.992 (7)**	**.997 (28)**	**.907 (3)**	**.899 (2)**	**.885 (2)**	.924 (28)	**.951 (1)**
$\mathcal{B}^{Comb}_{\tau,4}$	**.974 (6)**	.987 (27)	*.998 (14)*	1.004 (36)	*.910 (5)*	*.901 (4)*	*.887 (3)*	.924 (27)	*.953 (2)*
$\mathcal{B}^{Comb}_{\tau,5}$	*.981 (16)*	.999 (65)	*1.003 (24)*	1.009 (43)	*.920 (8)*	*.912 (5)*	*.894 (6)*	.927 (39)	*.961 (5)*
$\mathcal{B}^{Comb}_{\tau,6}$	.989 (38)	.999 (61)	1.015 (56)	1.006 (39)	*.919 (7)*	*.933 (10)*	*.911 (21)*	.930 (42)	**.968 (11)**
$\mathcal{B}^{Comb}_{\tau,7}$	.991 (48)	1.002 (78)	1.026 (79)	1.039 (84)	*.951 (20)*	*.944 (17)*	.914 (29)	.931 (44)	*.978 (24)*
$\mathcal{B}^{Comb}_{\tau,8}$	.991 (51)	1.002 (80)	1.025 (76)	1.037 (81)	*.948 (18)*	*.945 (22)*	.914 (30)	.932 (46)	*.978 (23)*
B^{RW}_{τ}	.981 (14)	1.002 (79)	1.017 (60)	1.009 (45)	*.908 (4)*	*.913 (6)*	*.899 (10)*	.944 (52)	*.964 (6)*
House price bubbles									
HP_{rec}									1.058 (154)
HP_{rol}	.988 (36)	1.005 (95)	**1.049 (95)**		**.951 (22)**	**.946 (23)**	.903 (14)	.921 (26)	.981 (33)
$PWY11^{i}_{rec}$	**.981 (17)**	**.998 (59)**			.958 (32)	.948 (26)	**.897 (8)**	**.906 (7)**	**.977 (19)**
$PWY11^{r}_{rec}$	.983 (22)	1.003 (88)			.976 (68)	.983 (76)	.923 (37)	.925 (32)	.990 (45)
$PWY11^{i}_{rol}$	.998 (88)					.998 (90)	.960 (89)	.970 (89)	1.014 (108)
$PWY11^{r}_{rol}$	.998 (85)					.985 (77)	.923 (39)	.926 (37)	1.014 (112)
$PSY15^{i}$	.985 (28)	1.003 (89)			.974 (67)	.981 (74)	.910 (20)	.917 (21)	.991 (47)
$PSY15^{r}$	.990 (43)				.986 (76)	.995 (86)	.947 (74)	.925 (31)	1.004 (80)
$\mathcal{B}^{Comb}_{\tau,1}$	**.983 (20)**	**1.001 (73)**			.983 (72)	.985 (78)	.940 (59)	.931 (45)	.993 (53)
$\mathcal{B}^{Comb}_{\tau,2}$	.986 (31)	1.003 (87)			.990 (82)		.945 (72)	.925 (34)	.998 (67)
$\mathcal{B}^{Comb}_{\tau,3}$	.987 (33)				.973 (64)	.987 (80)	.923 (38)	**.920 (24)**	.994 (58)
$\mathcal{B}^{Comb}_{\tau,4}$	.997 (79)					1.002 (99)	.963 (93)	.967 (83)	1.014 (111)
$\mathcal{B}^{Comb}_{\tau,5}$					.988 (80)	.992 (85)	.962 (92)	.971 (90)	1.011 (101)
B^{RW}_{τ}	.983 (21)	1.003 (86)	**1.032 (86)**	**1.056 (110)**	**.945 (16)**	**.942 (13)**	**.914 (28)**	.925 (35)	**.978 (21)**

Notes: The table shows the forecast performance the (augmented) AR models during NBER expansion periods. The second line shows the number of months defined as expansion periods. Only bubble indicators that rank among the best 100 predictors (or the best stock and house price bubble indicator at a given horizon) are shown for ease of presentation. See Table 3.4.1 for further notes.

Table 3.B.5: ARX forecasts: NBER recession periods

Horizon	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**	**Mean**
Obs.	34	34	34	34	34	34	34	34	
AR	13.897 (112)	12.295 (171)	11.624 (183)	11.754 (186)	1.961 (185)	9.636 (181)	7.452 (203)	5.835 (212)	1.432 (191)
DFM	.962	.937	.960	.937	.921	.957	.877	.912	.937
DFM ELA	.897	.845	.858	.919	.908	.921	.849	.762	.877
				Best non-bubble indicator					
SP500.P12	.868 (1)								.879 (16)
SP500.P12		.780 (1)							.879 (16)
SP500.P1			.678 (1)						.842 (7)
ENRG				.677 (1)					.739 (3)
OILAVG					.587 (1)				.738 (2)
OILAVG						.548 (1)			.738 (2)
METAL							.595 (1)		.780 (6)
METAL								.518 (1)	.780 (6)
				Stock price bubbles					
HP_{rec}	**1.000 (108)**	**.953 (60)**	**.896 (30)**	**.896 (30)**	.937 (93)		.921 (85)	**.880 (62)**	**.933 (51)**
HP_{rol}		.960 (68)	.905 (41)	.905 (35)	**.923 (47)**				.956 (107)
$PWY11^{i}_{rec}$									.965 (120)
$PWY11^{r}_{rec}$									.965 (121)
$PWY11^{i}_{rol}$			.931 (94)	.922 (46)	.931 (74)	.934 (97)			.951 (92)
$PWY11^{r}_{rol}$				.938 (96)	.933 (81)				.967 (123)
$PSY15^{i}$			.933 (96)	.923 (50)	.925 (51)	**.915 (60)**	**.901 (57)**	.905 (96)	.946 (76)
$PSY15^{r}$					.928 (61)	.925 (72)	.919 (80)	.903 (92)	.966 (122)
$\mathcal{B}^{Comb}_{\tau,1}$				.938 (100)				.901 (90)	.960 (111)
$\mathcal{B}^{Comb}_{\tau,2}$					.935 (85)	.929 (91)	**.911 (69)**	**.879 (59)**	.958 (109)
$\mathcal{B}^{Comb}_{\tau,3}$			.928 (85)	**.916 (37)**	.929 (64)				.950 (86)
$\mathcal{B}^{Comb}_{\tau,4}$			.928 (88)	.918 (40)	.931 (76)				.952 (97)
$\mathcal{B}^{Comb}_{\tau,5}$			.930 (91)	.921 (44)	.930 (69)	.934 (99)			.951 (95)
$\mathcal{B}^{Comb}_{\tau,6}$		**.974 (100)**	**.925 (75)**	.921 (43)	**.925 (49)**	**.921 (68)**	.916 (76)	.908 (99)	**.944 (73)**
$\mathcal{B}^{Comb}_{\tau,7}$			.927 (84)	.922 (48)	.928 (63)	.926 (76)	.920 (83)	.904 (93)	.946 (75)
$\mathcal{B}^{Comb}_{\tau,8}$			.928 (86)	.923 (52)	.930 (68)	.927 (86)	.923 (91)	.905 (95)	.947 (79)
B^{RW}_{τ}									.962 (113)
				House price bubbles					
HP_{rec}		**.972 (97)**	**.913 (55)**	.928 (70)					.953 (101)
HP_{rol}			.925 (74)	**.917 (38)**	**.923 (46)**	**.915 (59)**	**.899 (54)**	**.870 (53)**	**.940 (67)**
$PWY11^{i}_{rec}$			.933 (97)	.927 (65)	.935 (87)	.927 (87)	.918 (79)	.886 (68)	.951 (89)
$PWY11^{r}_{rec}$				.930 (80)	.940 (99)		.928 (99)	.893 (81)	.955 (106)
$PWY11^{i}_{rol}$			.927 (81)	.926 (62)	.930 (73)	.927 (79)			.950 (88)
$PWY11^{r}_{rol}$				.928 (71)	.935 (89)	.927 (78)	.913 (72)	.899 (89)	.951 (94)
$PSY15^{i}$			.932 (95)	.926 (60)	.933 (80)	.931 (94)	.918 (78)	.885 (67)	.950 (87)
$PSY15^{r}$	**1.012 (160)**	.973 (98)	.919 (66)	.932 (87)	.938 (96)	.929 (92)	.911 (68)	.902 (91)	.946 (77)
$\mathcal{B}^{Comb}_{\tau,1}$			**.926 (78)**	.929 (78)	.939 (98)		.914 (75)	**.884 (65)**	**.948 (80)**
$\mathcal{B}^{Comb}_{\tau,2}$			.934 (98)	.926 (64)	.937 (94)		.925 (96)	.893 (80)	.953 (100)
$\mathcal{B}^{Comb}_{\tau,3}$				.932 (83)	.938 (97)	.934 (100)	.921 (86)	.896 (84)	.955 (104)
$\mathcal{B}^{Comb}_{\tau,4}$			.928 (87)	.927 (67)	**.932 (77)**	.928 (88)			.953 (98)
$\mathcal{B}^{Comb}_{\tau,5}$			.925 (77)	**.924 (56)**	.929 (65)	**.925 (75)**			.949 (84)
B^{RW}_{τ}				.927 (68)	.934 (82)	.927 (80)	**.917 (77)**		.953 (99)

Notes: The table shows the forecast performance of selected predictors during NBER recession periods. The second line shows the number of months defined as recession periods. Only bubble indicators that rank among the best 100 predictors (or the best stock and house price bubble indicator at a given horizon) are shown for ease of presentation. See Table 3.B.4 for further notes.

Table 3.B.6: Forecast from augmented DFM with elastic-net preselection: NBER recession periods

Horizon	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**	**Mean**
DFM ELA	12.468 (26)	1.394 (28)	9.972 (5)	1.801 (5)	9.951 (10)	8.874 (25)	6.324 (20)	4.448 (12)	9.154 (19)
				Stock price bubbles					
HP_{rec}	**.981 (3)**	**.969 (2)**	**.984 (3)**	1.000 (6)					.993 (8)
HP_{rol}	.985 (5)	.976 (4)	.988 (4)	**.978 (1)**	**.969 (1)**	**.949 (1)**	**.963 (1)**	.989 (3)	**.975 (1)**
$PWY11^{i}_{rec}$	.985 (4)	.982 (5)							1.002 (21)
$PWY11^{r}_{rec}$									1.008 (25)
$PWY11^{i}_{rol}$									.996 (12)
$PWY11^{r}_{rol}$									1.036 (30)
$PSY15^{i}$	.987 (8)	.984 (8)				.978 (6)	.986 (5)	**.986 (2)**	.992 (6)
$PSY15^{r}$									1.044 (32)
$B^{Comb}_{\tau,1}$				**1.005 (9)**				**.939 (1)**	.997 (15)
$B^{Comb}_{\tau,2}$									1.014 (28)
$B^{Comb}_{\tau,3}$	.987 (7)	**.982 (6)**	1.001 (7)		.999 (6)	**.975 (3)**	**.982 (3)**	.998 (7)	**.992 (4)**
$B^{Comb}_{\tau,4}$	**.986 (6)**	.983 (7)	**1.001 (6)**		**.999 (5)**	.977 (4)	.984 (4)		.992 (5)
$B^{Comb}_{\tau,5}$		.986 (10)	1.002 (10)	1.006 (10)	1.000 (9)	.980 (8)	.987 (7)	.999 (10)	.993 (9)
$B^{Comb}_{\tau,6}$		.985 (9)			1.000 (8)	.978 (5)	.987 (6)	.990 (4)	.994 (10)
$B^{Comb}_{\tau,7}$						.982 (10)		.994 (5)	.996 (13)
$B^{Comb}_{\tau,8}$								.998 (8)	.998 (17)
B^{RW}_{τ}	.987 (10)		1.001 (8)				.989 (10)		.996 (11)
				House price bubbles					
HP_{rec}									1.041 (31)
HP_{rol}			1.002 (9)	.999 (4)	.993 (3)	.98 (9)	**.982 (2)**	**.999 (9)**	.992 (7)
$PWY11^{i}_{rec}$				1.005 (7)	.999 (7)				.997 (16)
$PWY11^{r}_{rec}$				1.005 (8)					1.001 (20)
$PWY11^{i}_{rol}$									1.006 (24)
$PWY11^{r}_{rol}$									1.023 (29)
$PSY15^{i}$									1.004 (22)
$PSY15^{r}$	**.975 (1)**	**.962 (1)**	**.969 (1)**	**.992 (3)**	**.992 (2)**	**.978 (7)**	.989 (8)		**.980 (2)**
$B^{Comb}_{\tau,1}$	**.981 (2)**	**.974 (3)**	**.972 (2)**	**.991 (2)**	**.998 (4)**	**.973 (2)**	**.989 (9)**	**.996 (6)**	**.983 (3)**
$B^{Comb}_{\tau,2}$									1.009 (26)
$B^{Comb}_{\tau,3}$									1.010 (27)
$B^{Comb}_{\tau,4}$									1.005 (23)
$B^{Comb}_{\tau,5}$									.999 (18)
B^{RW}_{τ}	.987 (9)								.996 (14)

Notes: The table shows the forecast performance of the (augmented) dynamic factor model (DFM) with elastic-net preselection during NBER recession periods. See Table 3.B.3 for further notes.

Table 3.B.7: Full sample parameter estimates for stock price bubble indicators

Horizon	**0**	**1**	**3**	**6**	**9**	**12**	**18**	**24**
HP_{rec}	.200 (4.32)	.190 (5.38)	.138 (4.47)	.110 (3.93)	.098 (3.79)	-.251 (.61)	-.191 (.54)	-.115 (.37)
HP_{rol}	.199 (4.30)	.187 (5.34)	.139 (4.50)	.110 (3.96)	.099 (3.84)	-1.219 (3.01)	-1.667 (4.72)	-1.639 (5.29)
$PWY11^{i}_{rec}$	.195 (4.22)	.186 (5.27)	.136 (4.41)	.107 (3.85)	.095 (3.71)	2.318 (3.04)	2.086 (3.14)	1.929 (3.34)
$PWY11^{r}_{rec}$	.198 (4.28)	.188 (5.32)	.136 (4.41)	.107 (3.84)	.095 (3.69)	2.045 (2.58)	1.663 (2.41)	1.227 (2.04)
$PWY11^{i}_{rol}$	.200 (4.33)	.190 (5.38)	.138 (4.46)	.109 (3.92)	.098 (3.78)	-1.113 (1.79)	-1.119 (2.07)	-1.093 (2.32)
$PWY11^{r}_{rol}$	.200 (4.32)	.189 (5.36)	.138 (4.48)	.110 (3.94)	.099 (3.81)	-.302 (.52)	-.497 (.99)	-.583 (1.33)
$PSY15^{i}$	.199 (4.30)	.189 (5.35)	.138 (4.46)	.109 (3.91)	.098 (3.77)	.343 (.71)	.203 (.48)	.089 (.24)
$PSY15^{r}$	.199 (4.31)	.189 (5.36)	.137 (4.46)	.109 (3.91)	.098 (3.77)	-.864 (1.56)	-.947 (1.96)	-.970 (2.31)
$B^{Comb}_{\tau,1}$	.200 (4.33)	.190 (5.41)	.139 (4.52)	.110 (4.01)	.099 (3.89)	-1.527 (4.09)	-1.615 (4.98)	-1.468 (5.17)
$B^{Comb}_{\tau,2}$	.197 (4.26)	.188 (5.31)	.137 (4.44)	.108 (3.88)	.097 (3.74)	.468 (1.13)	.284 (.78)	.307 (.97)
$B^{Comb}_{\tau,3}$	.200 (4.32)	.189 (5.36)	.138 (4.47)	.109 (3.92)	.098 (3.78)	-.020 (.04)	-.269 (.56)	-.377 (.90)
$B^{Comb}_{\tau,4}$	.199 (4.31)	.189 (5.36)	.138 (4.47)	.109 (3.92)	.098 (3.78)	-.081 (.14)	-.306 (.61)	-.438 (1.01)
$B^{Comb}_{\tau,5}$	.200 (4.33)	.189 (5.37)	.138 (4.47)	.110 (3.93)	.098 (3.79)	-.636 (1.04)	-.791 (1.49)	-.892 (1.93)
$B^{Comb}_{\tau,6}$	.198 (4.28)	.188 (5.32)	.137 (4.45)	.109 (3.89)	.097 (3.76)	.563 (.75)	.003 (.00)	-.277 (.49)
$B^{Comb}_{\tau,7}$	.199 (4.31)	.189 (5.36)	.138 (4.47)	.109 (3.93)	.098 (3.79)	.983 (1.07)	.853 (1.07)	.479 (.69)
$B^{Comb}_{\tau,8}$	.199 (4.29)	.188 (5.34)	.138 (4.46)	.109 (3.92)	.098 (3.78)	1.017 (.97)	1.151 (1.26)	1.031 (1.29)
B^{RW}_{τ}	.199 (4.31)	.189 (5.36)	.138 (4.47)	.109 (3.93)	.098 (3.79)	.110 (.21)	-.178 (.38)	-.317 (.78)

Notes: The table shows full-sample parameter estimates of the bubble indicators for different horizons h given in the top row. The parameters are obtained from estimating the augmented autoregressive model given in (3.2). Absolute values of the t-statistics are provided in parentheses.

CHAPTER 4

Monetary Policy and Mispricing in Stock Markets[1]

4.1 Introduction

The appropriate response of central banks to excessive asset price developments is subject of intense debate. While some researchers and policy makers advance the conventional view that central banks should focus solely on price and output stabilization (Bernanke and Gertler, 1999; Posen, 2006), others call for an active, "leaning against the wind" (LATW) monetary policy (Borio and Lowe, 2002; Cecchetti et al., 2002). This policy prescribes that central banks should monitor asset prices closely for the presence of bubbles, defined as periods when prices deviate excessively from their fundamentally justified value. If a bubble is detected, the central bank should raise the policy rate to attenuate asset mispricing and to contain the risks for financial and macroeconomic stability.

For an asset price targeting monetary policy to be feasible, however, two key requirements need to be fulfilled. First, a monetary policy tightening must not only lower asset prices, but specifically the possibly non-zero mispricing component in asset prices. Second, this response of the mispricing component needs to be suffi-

[1]This chapter is based on joint work with Kerstin Bernoth. We thank Klaus Adam, Refet Gürkaynak, Helmut Herwartz, Helmut Lütkepohl, as well as participants at the *IFABS 2016 Conference*, Barcelona; the *14th INFINITI Conference on International Finance*, Dublin; the *33rd International Symposium on Money, Banking and Finance*, Clermont-Ferrand; the *Conference on "Challenges of European Integration"* at University of Bonn, the 2015 DIW Berlin Macroeconometric Workshop, and seminar participants at Bundesbank, Norges Bank, University of Hamburg, University of Göttingen and Halle Institute for Economic Research (IWH) for helpful comments and suggestions.

ciently large such that a monetary intervention is not too costly in terms of inflation or output depression.[2] Hence, the LATW policy prescription builds strongly on the notion that an increase in interest rates reduces an asset price bubble substantially. Yet, this paradigm is challenged by Galí (2014), who argues that the bubble component in asset prices must grow in expectation at the rate of interest according to the theory on rational asset price bubbles. This argument receives empirical support from Galí and Gambetti (2015), who find that the bubble component increases following a monetary policy tightening. Raising the policy rate to combat an emerging asset price bubble would hence carry adverse effects and, in fact, exacerbate mispricing. Importantly, this also implies that a central bank faces a trade-off between stabilizing asset markets and inflation when an asset price boom is accompanied by a boom in real activity.[3]

In this paper, we challenge the theoretical predictions and empirical findings of Galí (2014) and Galí and Gambetti (2015) along three dimensions. First, we extend the theoretical analysis beyond the concept of rational bubbles to a more general notion of stock mispricing following Brunnermeier and Julliard (2008). More specifically, we allow for mispricing to arise (i) from a violation of the transversality condition; and/or (ii) from false expectations of irrational investors about the stock's underlying fundamentals, namely discounted future dividends and equity risk premia. While we do not attempt to empirically identify the source of mispricing, this framework carries important implications for the predicted response of stock mispricing to monetary policy shocks. Specifically, we emphasize that the ultimate effect of a monetary policy shock on the mispricing component in stock prices is ambiguous both in the short and in the long run and depends on how market participants form their expectations. Only if one assumes that all agents are rational, mispricing arises in the form of a rational bubble and one can remove the ambiguity about its response to obtain long-run bubble growth following a contractionary policy shock as in Galí (2014).

Second, we show that when the change in the required equity risk premium is not taken into account as in Galí and Gambetti (2015), any deviation from the funda-

[2]Furthermore, it is required that asset price bubbles can be detected in real-time. Promising real-time indicators have been proposed by Phillips et al. (2011, 2015) as discussed in Chapter 2.

[3]As argued in Chapter 3 this is likely the case since asset price booms alleviate borrowing constraints for households and firms, and by this spur consumption and investment (Kiyotaki and Moore, 1997; Bernanke et al., 1999).

mental stock price would be incorrectly attributed to mispricing. Our theoretical framework predicts stock prices to fall in response to a contractionary monetary policy shock not only due to the rational updating of agents' expectations about lower future dividends, higher real interest rates, but also due to a higher equity risk premium demanded by investors. All these factors must hence be accounted for before inference on the response of the mispricing component is possible. To obtain an unbiased estimate of the mispricing component, we therefore need measures for the expectations about both future dividends and the equity premium (Cochrane, 2011). Following a broad literature initiated by Campbell (1991), we obtain these expectations from forecasts of a stock pricing Vector Autoregressive (VAR) model. We then empirically decompose the stock price response to monetary policy shocks into its three economic sources: the response of the expected fundamental component related to the stream of future discounted dividends, the response of the expected equity premium, and the response of the mispricing component. By this, we also extend the work of Bernanke and Kuttner (2005), who decompose the stock price into a fundamental component and the equity risk premium. Allowing also for the possibility of stock mispricing, we find a quantitatively similar contribution of the fundamental component's response, but a smaller contribution of the response of the expected equity premium to the overall response of stock prices to a monetary policy shock.

Third, we challenge the empirical findings of Galí and Gambetti (2015) by employing a less restrictive strategy to identify monetary policy shocks in the data. The authors use a recursive identification scheme building on the assumption that monetary policy shocks move stock prices instantaneously but that the central bank does not respond within the quarter to idiosyncratic stock price movements. Only under this scheme they find that the mispricing component in stock prices increases following a contractionary monetary policy shock. However, the assumption that the U.S. FED does not respond contemporaneously to stock price shocks (including economic news captured by stock price movements) is disputable and is rejected by Lütkepohl and Netsunajev (2014), who test this assumption by exploiting heteroskedasticity in the data. Further empirical evidence indicates that the FED reacts to stock price innovations, at least in some periods (Rigobon and Sack, 2004; Bjørnland and Leitemo, 2009; and Furlanetto, 2011). Thus, we allow for two-way contempora-

neous responses between stock prices and the policy rate. We implement this by imposing both zero restrictions on impact and in the long run following Bjørnland and Leitemo (2009) and by augmenting these with sign restrictions employing the methodological advances by Arias et al. (2014) and Binning (2013). Similar to Galí and Gambetti (2015), we evaluate the response of the mispricing component in asset prices to monetary policy shocks not only in a VAR with constant coefficients but also in a time-varying coefficient (TVC) VAR following the specification of Primiceri (2005). A TVC-VAR has the advantage that we can take into account that the size and direction of the stock price response to monetary policy shocks might not be constant over time but depends on the size of the mispricing component.

We find that a monetary policy tightening lowers stock prices significantly. However, only about one third of the impact response in stock prices can be attributed to changes in expectations about future dividends, the real interest rate, and the equity premium. Hence, we conclude that the reduction in stock prices largely results from a decline in the mispricing component. We also find that the response of the mispricing component to monetary policy shocks is larger during periods of (excessive) stock price booms such as during the dot-com bubble of the 1990s. By this, we provide support to the arguments of the proponents of an active LATW monetary policy that a contractionary monetary policy may help to combat excessive asset price growth.[4] However, this comes at the cost of lower economic growth and a significant reduction in inflation. We conclude from this that a LATW monetary policy might not always be the recommended policy tool to combat excessive stock price developments, especially when excessive stock price booms arise in times of recession and low inflation. Nonetheless, our results provide comfort to monetary policy makers who want to contain inflationary risks during times when a an asset price boom is accompanied by a boom in real economic activity. In contrast to the implications of Galí (2014), we find that raising the policy rates to reduce inflation does not increase but also reduces mispricing on asset markets. Finally, our results are difficult to align with the theoretical predictions of the rational bubble framework, but are in accord with a concept of stock mispricing arising from false subjective beliefs of irrational investors.

[4]This finding is, of course, subject to the Lucas' critique (Lucas, 1976).

The remainder of this paper is structured as follows. Section 4.2 formally lays out our accounting framework to decompose stock prices into a fundamental component, the expected risk premium, and a mispricing component. We also discuss how and through which channels monetary policy affects these different stock price determinants. Section 4.3 outlines our empirical approach: We describe the TVC-VAR model and the identification strategy employed in our empirical analysis. Section 4.4 describes the data and our measures of expected dividends and the expected equity risk premium. We discuss our results in section 4.5. Section 4.6 concludes.

4.2 An accounting framework for asset prices

We employ a simple accounting framework based on the asset pricing equation to decompose the stock price into its fundamental component, an equity risk premium, and a mispricing component. This framework helps to guide our empirical impulse response accounting strategy and lays out the theoretically predicted impact of a monetary policy shock on the stock prices and its components. These theoretical predictions will then guide our identification strategy to disentangle monetary policy shocks from other structural shocks. Our framework differs from Galí (2014) and Galí and Gambetti (2015) in two ways: First, we relax the assumption that the expected stock price return (the discount factor) equals the risk free return. Instead we allow for the presence of a (time-varying) expected equity risk premium. Second, we generalize their concept of the mispricing component potentially present in stock prices. The authors discuss theoretical predictions for the response of this mispricing component to monetary policy shocks under the assumption that it takes the form of a rational bubble when all investors are fully rational. We abstract from this narrow definition by following the argumentation of Campbell and Vuolteenaho (2004) and Brunnermeier and Julliard (2008) that stock mispricing may also occur due to false subjective believes of irrational investors on future fundamentals and risk premia.

We begin with the standard asset pricing equation. Let P_t and D_t be a stock's market price and dividend at the end of period t or accrued over period t, respectively. The net return on the stock between date t and $t+1$ is then given by:

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} - 1 \tag{4.1}$$

One can then show that the stock price is a function of future dividends D_{t+i}, the required net stock return R_{t+i}, and a terminal value:

$$P_t = \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i} \frac{1}{1+R_{t+j}} \right) D_{t+i} + \lim_{T\to\infty} \left(\prod_{j=1}^{T} \frac{1}{1+R_{t+j}} \right) P_{t+T}. \tag{4.2}$$

In log-linear form (less a constant) the stock pricing equation can then be written as:[5]

$$p_t = \sum_{i=1}^{\infty} \rho^{i-1} \left[(1-\rho) d_{t+i} - r_{t+i} \right] + \lim_{T\to\infty} \left(\rho^T p_{t+T} \right), \tag{4.3}$$

where logs of variables are denoted by lowercase letters, $r_{t+1} = \log(1+R_{t+1})$ and ρ is a parameter of the linearization defined as $\rho \equiv 1/\left(1+\exp(\overline{d-p})\right)$, where $\left(\overline{d-p}\right)$ is the long-run average log dividend-price ratio (such that $0 < \rho < 1$) (Campbell et al., 1996).[6]

We introduce an equity premium to the pricing equation by deducting the real risk free rate, r_t^f, measured by the real return on a long-term government bond, from both dividends and the required stock return. Hence, we can rewrite (4.3) in terms of excess dividends, $d_t^e = (1-\rho)d_t - r_t^f$, and excess returns (the equity risk premium) that compensate investors for holding risky equity instead of alternative safe investments, such as government bonds $r_t^e = r_t - r_t^f$:

$$p_t = \sum_{i=1}^{\infty} \rho^{i-1} d_{t+i}^e - \sum_{i=1}^{\infty} \rho^{i-1} r_{t+i}^e + \lim_{T\to\infty} \left(\rho^T p_{t+T} \right) \tag{4.4}$$

Thus, the stock price reflects the discounted value of future excess dividends, an equity premium measured by the discounted value of future excess stock returns, and a terminal value.

[5] A derivation of equation (4.3) is shown in Appendix 4.A.

[6] Campbell et al. (1996) shows with U.S. data that the average dividend-price ratio has been about 4% annually, implying $\rho \approx 0.99$ when applied to quarterly data.

4.2.1 Expectations and asset (mis)pricing

Following equation (4.4), today's stock price depends on future realizations of fundamentals and of the stock price itself. This means that agents' expectations about the future will determine the current stock price. Following Campbell and Vuolteenaho (2004) and Brunnermeier and Julliard (2008), we allow for the possibility that some investors are irrational, and hold subjective expectations at time t, denoted by $\tilde{E}_t$, about future realizations of d_t^e and r_t^e. These expectations obtained under the subjective probability measure may deviate from objective expectations E_t that are consistent with the rational processing of objective data.[7] Nevertheless, both expectations are conditioned on the same information set Ω_t available at time t, as indicated by the subscript on E and $\tilde{E}$. Hence, taking objective and subjective expectations of (4.4) yields

$$p_t = E_t\left[\sum_{i=1}^{\infty}\rho^{i-1}d_{t+i}^e\right] - E_t\left[\sum_{i=1}^{\infty}\rho^{i-1}r_{t+i}^e\right] + E_t\left[\lim_{T\to\infty}\left(\rho^T p_{t+T}\right)\right] \tag{4.5}$$

$$= \tilde{E}_t\left[\sum_{i=1}^{\infty}\rho^{i-1}d_{t+i}^e\right] - \tilde{E}_t\left[\sum_{i=1}^{\infty}\rho^{i-1}r_{t+i}^e\right] + \tilde{E}_t\left[\lim_{T\to\infty}\left(\rho^T p_{t+T}\right)\right], \tag{4.6}$$

respectively. The equality of equations (4.5) and (4.6) holds since the current observed stock price must reflect both rational and irrational investors' expectations about future excess dividends and equity premia for them to participate in the market. However, these expectations about the paths of both variables might differ between both types of investors. This means, for example, that irrational investors who require a high risk premium, $\tilde{E}_t\left[r_{t+i}^e\right]$, also expect to receive higher excess dividends in the future, $\tilde{E}_t\left[d_{t+i}^e\right]$. Yet, when irrational investors are present, the observed price p_t may deviate from its true fundamental value. This is observed by rational investors who, on average, form correct expectations about future fundamentals $E_t\left[d_{t+i}^e\right]$ and therefore adjust their required future equity risk premium $E_t\left[r_{t+i}^e\right]$ such that their stock holdings are in equilibrium. This change in $E_t\left[r_{t+i}^e\right]$ then ensures that (4.5) and (4.6) hold at the observed price level (Brunnermeier and Julliard, 2008).

[7] For a further elaboration on this concept, see Manski (2004) and Brunnermeier and Parker (2005).

By adding and subtracting $\left(E_t\left[\sum_{i=1}^{\infty}\rho^{i-1}d^e_{t+i}\right]-E_t\left[\sum_{i=1}^{\infty}\rho^{i-1}r^e_{t+i}\right]\right)$ from (4.6), we observe that the stock price p_t can be written as the sum of three components: a fundamental component, p^F_t, measured by the discounted value of expected future excess dividends; an equity premium, ep_t, measured by the discounted value of expected future excess stock returns; and a mispricing component, ψ_t:

$$p_t = \underbrace{E_t\left[\sum_{i=1}^{\infty}\rho^{i-1}d^e_{t+i}\right]}_{p^F_t} - \underbrace{E_t\left[\sum_{i=1}^{\infty}\rho^{i-1}r^e_{t+i}\right]}_{ep_t} + \psi_t, \tag{4.7}$$

where ψ_t is defined as:

$$\psi_t = (\tilde{E}_t - E_t)\left[\sum_{i=1}^{\infty}\rho^{i-1}d^e_{t+i}\right] - (\tilde{E}_t - E_t)\left[\sum_{i=1}^{\infty}\rho^{i-1}r^e_{t+i}\right] + \tilde{E}_t\left[\lim_{T\to\infty}\left(\rho^T p_{t+T}\right)\right]. \tag{4.8}$$

Thus, the observed stock price p_t deviates from its objectively justified fundamental value when (i) irrational investors are present, whose subjective expectations deviate from objective expectations of rational investors; or (ii) the transversality condition under the subjective measure does not hold, i.e. $\tilde{E}_t\left[\lim_{T\to\infty}\left(\rho^T p_{t+T}\right)\right] \neq 0$. If one assumes, however, that all agents form expectations under the objective probability measure, mispricing may only result from a violation of the transversality condition under objective expectations, i.e. $\psi_t = E_t\left[\lim_{T\to\infty}\left(\rho^T p_{t+T}\right)\right]$. Thus, only if all investors are of the rational type and ψ_t is non-zero, mispricing occurs due to the existence of a rational bubble and investors are fully aware of it.[8] Without these restrictive assumptions, however, mispricing can result from false subjective expectations of investors. This is in line with Adam et al. (2015), who also show that subjective belief dynamics can temporarily delink stock prices from their fundamental value and give rise to asset price booms.[9] As such, the concept presented here is less restrictive than the one described by Galí (2014) and Galí and Gambetti (2015), who explain mispricing entirely with the presence of a rational bubble.

[8] For example, this can be explained in the context of overlapping generations models (see the survey of Stiglitz (1990) for a larger discussion) or in the context of intrinsic bubbles, as introduced by Froot and Obstfeld (1991).

[9] For a further extensive survey on the literature relating speculative behavior to irrational and behavioral factors, see Scherbina and Schlusche (2014).

4.2.2 Effects of monetary policy on stock prices

Our accounting framework illustrates that monetary policy shocks may affect stock prices via all its three components. To guide our empirical analysis, we discuss these three sources in turn.

First, economic theory points to a decrease of the fundamental component of stock prices in response to a contractionary monetary policy shock for two reasons: First, a monetary tightening decreases future economic growth and, by this, also firms' profits and future dividend pay-outs. Moreover, since inflation is expected to decrease following the contractionary policy shock and nominal interest rates increase, real risk-free rates also rise (Bernanke and Kuttner, 2005). As a result, the fundamental component is predicted to fall in response to a contractionary monetary policy shock:[10]

$$\frac{\partial p^F_{t+k}}{\partial \varepsilon^m_t} = \left[\sum_{i=1}^{\infty} \rho^{i-1} \frac{\partial E_t(d^e_{t+k+i})}{\partial \varepsilon^m_t} \right] < 0 \tag{4.9}$$

Second, the expected equity premium is likely to rise following an exogenous monetary tightening for two reasons. First, there is evidence for a (non-linear) financial accelerator mechanism of monetary policy via balance sheet effects of firms and banks, which implies that a monetary tightening is associated with higher borrowing costs and reduced loan supply. In effect, this increases the risk of lower future firm profits. Importantly, the increase in risk can be expected to be larger during times of already tight monetary policy when loan supply is already low and firms' financial health is poor, which calls for a modeling strategy that accounts for this time-variation (Patelis, 1997). Second, as outlined by Gust and López-Salido (2014), a contractionary monetary policy may also reduce participation in risky asset markets and thereby decrease market liquidity and risk sharing, which would result in a

[10] An opposing prediction could be obtained if the monetary policy surprise provides new information to market participants who are less informed about future output and inflation than the central bank. A monetary policy tightening may then signal higher expected future economic growth and, thus, lead to an upward revision of market participants' expectations about future excess dividends. However, we assume that market participants and the central bank share the same information set. Second, by restricting excess dividends to fall following a contractionary monetary policy shock, we also offer fundamental variables the best chance to explain a possible decrease in stock prices. Our empirical results are quantitatively robust to when we allow for a positive response of expected excess dividends.

higher equity risk premium required by investors today. Hence, an exogenous monetary tightening can be expected to raise the equity premium that rational investors require to be compensated for the increase in risk:

$$\frac{\partial ep_{t+k}}{\partial \varepsilon_t^m} = \left[\sum_{i=1}^{\infty} \rho^{i-1} \frac{\partial E_t(r^e_{t+k+i})}{\partial \varepsilon_t^m} \right] > 0. \tag{4.10}$$

Since the expected equity risk premium enters the stock pricing equation negatively, this lowers the current stock price further. Therefore, in absence of a mispricing component, a contractionary monetary policy shock would induce an immediate fall in stock prices through both the fundamental and the risk premium component.

In contrast to the responses of these components, the response of the mispricing component ψ_t is open to debate. The conventional view of LATW policy proponents builds on the notion that (excessive) stock prices typically fall in response to a contractionary monetary policy shock. In a similar vein, a monetary loosening should therefore propagate stock overpricing. This claim finds empirical support from e.g. Borio and Lowe (2002), who argue that accommodative monetary policy may, in a low-inflationary environment, stimulate asset price bubbles. Similarly, Bordo and Landon-Lane (2013) find that several measures of loose monetary policy can be consistently related to periods that are demarcated as asset price booms, even when controlling for other explanatory variables such as credit and current account imbalances. However, these papers are silent about whether an observed asset price boom is driven by stronger fundamentals and a lower equity premium, or by mispricing, and therefore do not allow for drawing any conclusions on the justification for a LATW monetary policy. Here, the findings of Adam et al. (2015) provide some first tentative evidence that a contractionary monetary policy shock may help to attenuate excessive stock price behavior. The authors find that about two-thirds of the fluctuations in U.S. stock prices are not due to fundamental factors, but result from self-reinforcing beliefs in the sense that agents become more pessimistic about future capital gains whenever they are negatively surprised by past capital gains, and vice versa. A negative capital gains surprise in the previous period then increases pessimism about future capital gains and leads to an asset price bust. Thus, this result indicates that an unexpected decrease in capital gains

in response to a contractionary monetary policy shock would lead to a downward correction of the excessive stock price component.

However, the view that contractionary monetary policy helps to reduce excessive stock prices is questioned by Galí (2014). If all investors are fully rational and form objective expectations about the future, mispricing may only occur in the form of a rational bubble. In this case, the bubble component must grow at the required rate of return on stocks in expectation.[11] Thus, since both the risk-free interest rate and the expected equity premium are predicted to increase in response to a contractionary policy shock, an increase in the policy rate will also raise the expected long-run growth rate of the bubble component. This holds, however, only in expectation for each period *after* the policy shock. On impact, the response of a rational bubble component is, in fact, indeterminate.[12] However, Galí (2014) disregards that mispricing may also occur from the presence of irrational investors.

Thus, the impact of monetary policy shocks on the stock mispricing component depends on how investors form their expectations and is therefore not only indeterminate on impact but also in the long run, i.e.:

$$\frac{\partial \psi_{t+k}}{\partial \varepsilon_t^m} = \left[\sum_{i=1}^{k} \rho^{i-1}(\tilde{E}_t - E_t)\frac{\partial d_{t+i}^e}{\partial \varepsilon_t^m}\right] - \left[\sum_{i=1}^{k} \rho^{i-1}(\tilde{E}_t - E_t)\frac{\partial r_{t+i}^e}{\partial \varepsilon_t^m}\right] + \tilde{E}_t\left[\frac{\partial \rho^T p_{t+k}}{\partial \varepsilon_t^m}\right] = ? \tag{4.11}$$

To summarize, the responses of the fundamental component and of the equity risk premium to a contractionary monetary policy shock predict an immediate fall in the stock price. In contrast, the direction of the mispricing response is ambiguous. As a result, the total response of stock prices to a monetary policy shock ultimately depends on the nature and the size of the mispricing component relative to the stock's fundamentals. Since this relative size might vary over time, we need to opt for a flexible modeling strategy in our empirical analysis of monetary policy effects on stock prices.

[11] See Chapter 2 for a formal discussion of this result.
[12] See Galí (2014) for a discussion of this issue.

4.3 Empirical model and identification

To shed light on the direction and quantitative importance of the mispricing response to monetary policy shocks, we provide evidence from a standard monetary VAR model used to identify monetary policy shocks which we augment by stock prices, expected excess dividends, and the expected equity premium. We provide evidence from both a constant and a time-varying coefficient (TVC) VAR. The motivation for the latter is threefold. First, and most importantly, time-variation in the response of stock prices to monetary policy shocks may arise from the presence of the mispricing component. When mispricing is small, the fundamental and risk premium response are likely to determine the overall stock price response. In times of pronounced mispricing, however, its response is likely to assert a stronger influence on the overall stock price response. Second, as outlined in the previous section, we also suspect non-linearities in the relationship between measures for the business cycle, the policy rate, and our measure for the expected equity risk premium following the discussion in Patelis (1997) and Gust and López-Salido (2014). Therefore, independent of strong empirical evidence for time-varying coefficients, our model framework demands such an assessment. Finally, the TVC-VAR has the general advantage that it allows for accounting for structural breaks and smooth structural change in the model coefficients. Since the constant coefficient VAR (C-VAR) is a limiting case of the TVC-VAR, as described below, we will only outline the specifications of the TVC-VAR.

4.3.1 Time-varying coefficient VAR

Our reduced form TVC-VAR follows closely the specification developed by Primiceri (2005) and also employed by Galí and Gambetti (2015).[13] The dynamic relations between the variables are described by the measurement equation:

$$Y_t = \Theta_{0,t} + \Theta_{1,t} Y_{t-1} + \cdots + \Theta_{p,t} Y_{t-p} + u_t \tag{4.12}$$

[13] We use and adapt the Matlab code for the estimation of the TVC-VAR and the subsequent impulse response analysis provided by Galí and Gambetti (2015) at `https://www.aeaweb.org/articles?id=10.1257/mac.20140003`. For the reduced-form estimation of the C-VAR, we adapt the code provided by Gary Koop at `https://sites.google.com/site/garykoop/home/computer-code-2`. Any errors remain our own.

where Y_t is the $K \times 1$ vector of endogenous variables. The variables included in the VAR are, first, a measure of output growth y_t, inflation π_t, and the policy variable i_t (federal funds rate) as motivated by the standard class of New-Keynesian models used to identify the Taylor-rule type monetary policy shocks.[14] Following our partial equilibrium asset pricing model, we further add the growth rate in real stock prices p_t, and measures for expected excess dividend $E_t[d^e_{t+k}]$ and the expected equity risk premium $E_t[r^e_{t+k}]$ to the set of variables.[15] We specify the model with $p = 3$ lags of endogenous variables.[16] The reduced-form residuals u_t are assumed to be independently and normally distributed with variance-covariance matrix H_t, i.e. $u_t \sim N(0, H_t)$. The time-varying coefficients in $\Theta_{i,t}, i = 0, 1, \ldots, p$, evolve according to a driftless random walk. This is given by the state equation on the joint $m \times 1$ vector $\theta_t = vec([\Theta_{0,t}, \Theta_{1,t}, \ldots, \Theta_{p,t}])$ with $m = K(Kp + 1)$:

$$\theta_t = \theta_{t-1} + \eta_t. \tag{4.13}$$

Here, $\eta_t \sim N(0, Q)$ is a random vector, independent of u_t for all periods t and s.

The model allows for heteroskedasticity in the reduced-form residuals. To model changes in volatility, a triangular decomposition of the covariance matrix H_t of u_t is used, given by:

$$H_t = A_t^{-1} \Sigma_t \Sigma_t' (A_t^{-1})', \tag{4.14}$$

where Σ_t is a diagonal matrix with elements $\sigma_{i,t}$ for $i = 1, \ldots, K$ and A_t is lower triangular with diagonal elements equal to one.

The dynamics of covariances are governed by:

$$\alpha_t = \alpha_{t-1} + \omega_t, \tag{4.15}$$

[14] The responses of asset prices and the underlying fundamentals are robust with respect to including variables typically employed to capture expected inflation, such as commodity prices and non-energy commodity prices. Therefore, these are excluded in order to reduce the computational burden.

[15] In the following, the variables are ordered as $Y_t = [y_t, \pi_t, E_t[d^e_{t+k}], E_t[r^e_{t+k}], i_t, p_t]'$.

[16] The findings are robust to varying lag lengths of $p = 4$ lags as employed by Galí and Gambetti (2015).

where $\alpha_t = [a_{21,t}, \ldots, a_{K(K-1),t}]'$ captures the non-zero and non-unity elements of A_t and $\omega_t \sim N(0, W)$, with W being block-diagonal.[17] Finally, the dynamics of Σ_t are modeled according to a stochastic volatility framework, such that:

$$\log(\sigma_{i,t}) = \log(\sigma_{i,t-1}) + \nu_{i,t}, \tag{4.16}$$

with the vector $\nu_t = [\nu_{1,t}, \ldots, \nu_{K,t}]'$ being distributed as $\nu_t \sim N(0, V)$, where V is diagonal. The TVC-VAR nests the VAR with constant coefficients if the variance matrices Q, V, and W shrink to zero. The choice of prior distributions is briefly specified in Appendix 4.B and follows Galí and Gambetti (2015).

Finally, our core interest lies in the structural shocks ε_t and their effects on our endogenous variables y_t. We adopt the common assumption that the reduced-form residuals u_t are a linear transformation of the underlying structural shocks ε_t given by $u_t = S_t\varepsilon_t$, where S_t is the time-varying impact matrix. The structural shocks fulfill the common requirements $E[\varepsilon_t\varepsilon_t'] = I_K$ and $E[\varepsilon_t\varepsilon_{t-j}'] = 0$ for all t and $j \neq 0$, such that S_t satisfies $S_tS_t' = H_t$.[18] We identify S_t as outlined in the next subsection. With the identified S_t, we then obtain local approximations of the impulse responses to the period t structural shocks ε_t as Galí and Gambetti (2015).[19] By simple accounting, the response of the mispricing component can then be obtained by deducting the implied response paths of the expected fundamental component and the expected equity premium from the observed total response of asset prices following equation (4.7) similar to Galí and Gambetti (2015).

[17] Assuming block-diagonality implies that covariances across equations evolve independently from each other. This simplifies inference and drastically increases the efficiency of the estimation algorithm (Primiceri, 2005).

[18] All random vectors ε_t, η_t, ν_t and ω_t are assumed to be independent of each other. This restriction reduces the number of coefficients that need to be estimated and allows a structural analysis of the shocks ε_t that would be precluded if block elements in the covariance matrix of the vector $[\varepsilon_t, \eta_t, \nu_t, \omega_t]'$ were non-zero (Primiceri, 2005).

[19] As common in the literature (cf. Primiceri, 2005, Galí and Gambetti, 2015, or Prieto et al., 2016), we refrain from computing generalized impulse response functions (GIRFs) to the structural shocks in our TVC-VAR specification. We note that this is a necessary simplification, as computing GIRFs increases the computational burden of the empirical analysis. See Baumeister and Peersman (2013) for an outline how to implement the computation of GIRFs in TVC-VARs.

4.3.2 Identification via sign restrictions

An important conclusion that can be drawn from Galí and Gambetti (2015) is that their empirical results are not robust to the assumptions used to identify the impact matrix S_t and the structural shocks ε_t. Specifically, the authors obtain diametrically opposing results on the effect of monetary policy shocks on stock prices depending on whether they assume that the central bank responds within the same quarter to stock price surprises (by a calibrated coefficient) or not at all (as in their benchmark recursive Cholesky system). Allowing for contemporaneous responses of stock prices to monetary policy shocks, and of the interest rate to stock price shocks is hence crucial for our analysis. This simultaneity issue in identifying monetary policy shocks in the presence of financial variables is commonly known in the literature. Bjørnland and Leitemo (2009) provide a solution to this issue by augmenting an otherwise recursively identified system by one additional long-run restriction, imposing that monetary policy shocks carry no long-run effects on real stock prices according to the long-run neutrality of money. Hence, their system is just-identified and allows to freely estimate the impact response of monetary policy to stock price shocks. However, for our framework with three financial and forward-looking variables, we require an even less restrictive identification scheme with additional non-zero impact responses. Therefore, we need to relax the impact restrictions of the recursive identification scheme further, and hence we only partially identify the monetary policy shock. To then narrow our set of admissible impulse responses we augment the zero impact and long-run restrictions by imposing as many additional restrictions as we can defend from the theoretical considerations discussed in section 4.2. This can be implemented by employing sign restrictions on the impact responses (Fry and Pagan, 2011).[20] Thus, we allow for non-zero impact responses of stock prices, the policy rate, and both expectation measures to all shocks. By this we nest all alternative calibrations of the contemporaneous response of monetary policy to stock prices shocks of Galí and Gambetti (2015). Work by Arias et al. (2014) and Binning (2013) allows us to combine sign restrictions with zero impact and long-run restrictions.[21]

[20] An alternative could be to employ the external instruments approach as in Gertler and Karadi (2015). This approach is, however, not yet established for TVC-VARs.

[21] A limitation to our approach may be that our prior assumptions on the decomposition and time evolution of the variance-covariance matrix of residuals (equations (4.14)-(4.16)) is not

Table 4.3.1: Identifying impact restrictions

	Shock			
Variable	ε_t^d	ε_t^s	ε_t^m	ε_t^p
y_t	+	−	0	0
π_t	+	+	0	0
$E_t[d_{t+k}^e]$	?	?	−	?
$E_t[r_{t+k}^e]$	?	?	+	?
i_t	+	?	+	?
p_t	+	?	?	+

Notes: The table shows the sign restrictions on the impact responses of the variables in the benchmark VAR to a monetary policy shock ε_t^m, a stock price shock ε_t^p, as well as to a demand (ε_t^d) and a supply (ε_t^s) shock. The monetary policy shock is further restricted to have no long-run effect on the real variables y_t and p_t. Entries denoted with by a "?" are left unconstrained.

Table 4.3.1 summarizes our identifying assumptions on the impact responses of the variables in the VAR to four identified shocks. We assume that output and inflation do not respond within the same quarter to a (contractionary) *monetary policy shock* (ε_t^m). This assumption features strongly in the literature on identifying monetary policy shocks and mirrors the restrictions implied in the Cholesky decomposition (cf. Christiano et al., 2005), which are largely motivated by the perception that the transmission of monetary policy interventions to the real economy is only effective with a considerable lag (Friedman, 1968).[22] In contrast, all expectations and financial variables are allowed to respond immediately. The imposed signs are motivated by the theoretical predictions developed in section 4.2: Expected excess dividends decrease on impact while the expected equity premium increases.[23] Stock prices are also allowed to respond, yet the sign of the response is ambiguous, allowing for a strong positive response of the mispricing component.

fully agnostic and thus may fail to identify all admissible structural models. A solution to this limitation is unfortunately outside the scope of this paper. We thank Prof. Rubio-Ramírez for helpful consultations on this matter.

[22] For a robustness check, we relax the zero impact restrictions on ε_t^m, replacing them with sign restrictions that allow the contractionary monetary policy shock (both conventional and unconventional) to lower both output and inflation on impact (Baumeister and Benati, 2013). The reduction of output growth detected in this case is stronger, while the responses of stock prices and its components remain unaffected.

[23] However, for a robustness check, we also allow for the possibility that market participants revise their expectations about excess dividends upwards following the signal of a contractionary monetary policy shock. In this case, we leave the response of expected excess dividends unconstrained.

Following Bjørnland and Leitemo (2009), we interpret a *stock price shock* (ε_t^p) either as a non-fundamental, idiosyncratic shock motivated by speculative behavior, or as news about future real economic fundamentals. Therefore, expectations and financial variables may respond immediately to the arrival of new fundamental information or simply due to portfolio re-balancing in view of higher realized capital gains. The sign of their responses is left unconstrained with the obvious exception of stock prices. Similar to the case of the monetary policy shock, we assume that a stock price shock has no immediate effect on the slow moving variables output and inflation. This identification, therefore, distinguishes idiosyncratic stock price shocks from general demand or supply shocks. This is at the heart of the discussion regarding how central banks should react to stock price shocks that are not immediately related to their primary targets of inflation and output, but may only affect those variables in the longer run. Yet, this leaves us with the simultaneity problem in distinguishing stock price and monetary policy shocks, as they are not uniquely identified so far. We solve this issue by imposing additional long-run restrictions on the monetary policy shock following Bjørnland and Leitemo (2009). Specifically, we assume that monetary policy does not have any permanent effect on the real variables output and stock prices by the long-run neutrality of money.

Finally, we attempt to identify general aggregate *demand* (ε_t^d) and *supply shocks* (ε_t^s). Even though we are not primarily interested in those shocks, imposing as many restrictions as possible helps to close in on the shocks of interest and discard implausible models (cf. Canova and Gambetti, 2009 and Kilian and Murphy, 2012). The imposed impact restrictions displayed in the table are standard in the literature and follow Baumeister and Benati (2013). Following the previous discussion on the nature of a stock price shock, a general demand or supply shock is hence characterized by immediate responses of output and inflation. We do not constrain the unidentified fifth and sixth shock of the model, but assure that all shocks provide distinct sign patterns.

4.4 Data

Our analysis builds on sample information from 1962Q1 to 2014Q4. In our benchmark analysis we do not exclude the financial crisis period but rely on our TVC model to account for possible structural breaks and smooth structural changes in

the parameters. The variables included in the VAR are real GDP, the GDP deflator (both in log-growth rates), the effective federal funds rate and the growth rate of the real S&P 500 index.[24] All series are deflated by the U.S. Consumer Price Index for All Urban Consumers. Further, we add the two measures for *expected* excess dividends, $E_t\left[d^e_{t+k}\right]$, and and *expected* excess return on stocks, $E_t\left[r^e_{t+k}\right]$, (the expected equity risk premium). These variables are, however, not observable.

To obtain a proxy for both expectations, we therefore follow the broad literature initiated by Campbell (1991), which is also employed by Bernanke and Kuttner (2005) and Brunnermeier and Julliard (2008), and use an auxiliary VAR model to forecast the objective expectations of d^e_{t+k} and r^e_{t+k} at each margin t. The forecast $\hat{d}^e_{t+k}$ and $\hat{r}^e_{t+k}$ from the VAR can then be interpreted as the objective expectations $E_t\left[d^e_{t+k}\right]$ and $E_t\left[r^e_{t+k}\right]$ consistent with the rational processing of data in line with our model framework. The variables included in the auxiliary VAR model are *realized* log real dividends and the *realized* quarterly real stock return r_t, and the real risk free rate measured by the ten-year U.S. government bond rate. From the forecasted series of these variables, we obtain our measures for expected excess dividends and the equity risk premium. Additionally, we add the log price-earnings ratio, $p_t - e_t$, and the BAA-AAA corporate bond spread, r^b_t, as important predictors for future stock returns and equity risk (Campbell et al., 2013).[25] Here, the corporate bond spread serves as an additional measure for expected default risk in the economy, which is likely to be correlated with the expected equity risk premium.[26] The lag length of the VAR is set to one and forecasts are obtained from a VAR estimated on the full sample in line with the literature (Campbell, 1991; Bernanke and Kuttner, 2005; Brunnermeier and Julliard, 2008; Campbell et al., 2013). Finally, the forecast horizon is set to ten years to relate expected excess equity returns to the yields on

[24] Data is obtained from Federal Reserve Economic Data (FRED) of the Federal Reserve Bank of St. Louis, available at `http://research.stlouisfed.org/fred2/`, and from Shiller (2005), available at `http://www.econ.yale.edu/~shiller/data.htm`.

[25] Earnings e_t are the ten-year moving average of quarterly earnings from Shiller (2005).

[26] The forecasts are robust to additional variables frequently used in the literature, such as the term spread measured by the difference between the ten-year U.S. treasury bond and the three-month treasury bill yield (Campbell et al., 2010) and a measure for expected market volatility obtained from the quarterly variance of daily returns (Campbell et al., 2012).

the long-maturity government bond. Ten year forecasts are obtained sequentially for each quarter.[27]

Figure 4.4.1 shows the ten year forecasts of excess dividends and equity returns over time, which we interpret as the rational expectations at each point in time. In line with evidence in Claus and Thomas (2001), we find an expected equity premium of around three percent, which fluctuates, however, quite substantially over time. We find an increase in the expected premium simultaneous with the rise in real interest rates during the Volcker-disinflation period until about 1984. Subsequently, during the Great Moderation, the expected equity premium decreased back to the level of the early 1960s. From 1995 onwards, the expected equity risk premium declined further, contemporaneous to the sharp run-up in stock prices during the dot-com bubble. During the collapse of the bubble in the early 2000s, the expected equity premium jumped up again. In fact, the initial rise in the expected equity premium precedes the dot-com crash in 2000Q1. Finally, the expected equity premium spiked once more, but this time subsequently to the collapse of the housing bubble.

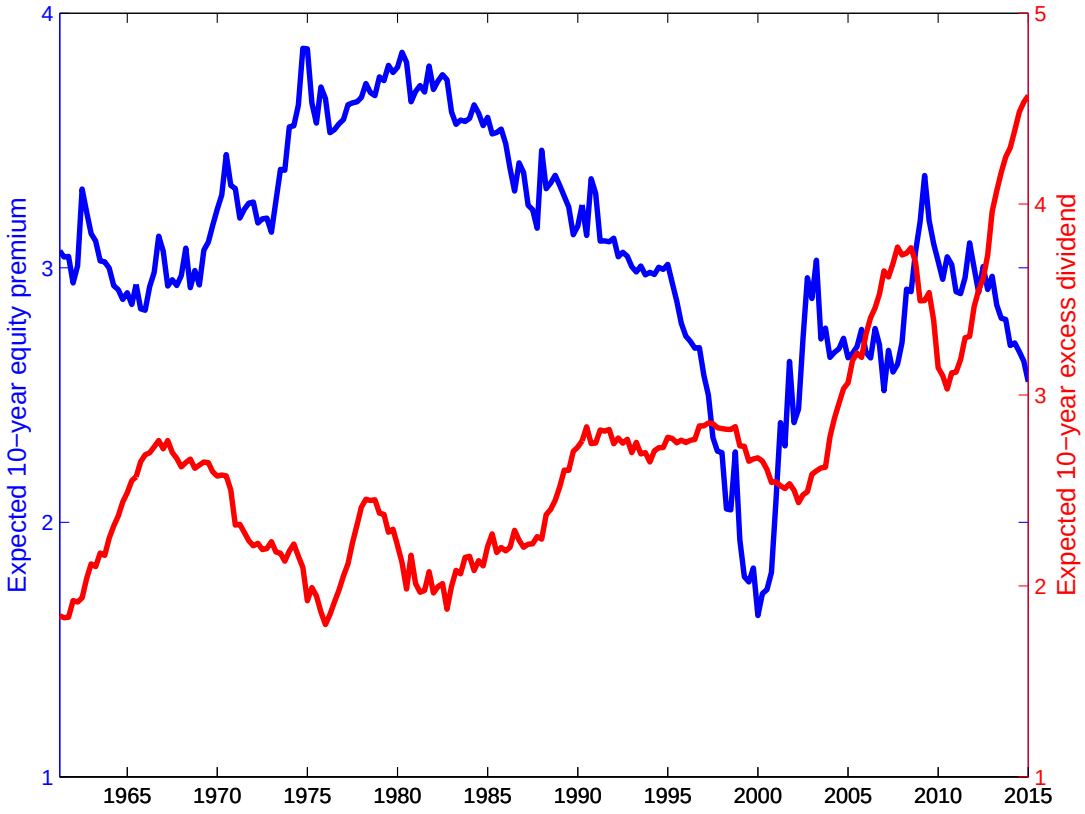
In contrast to the expected equity premium, expectations on excess dividends are relatively stable until about 2003. Thereafter they increased strongly with the exception of the period around the global financial crisis in 2008. Since the expected excess dividends series is found to be non-stationary, this variable enters our (TVC-)VAR in log-differences.

As a sensitivity check, we compare our measure of the expected equity risk premium with two alternative proxies of expected default or equity risk commonly used in the literature. The first is the BAA-AAA spread, which measures the expected default risk premium in the corporate bond market and should therefore be correlated with expected risks on equity. The second measure is derived from the Gordon (1962) dividend discount model (DDM), which suggests to infer the implied required rate of return on stocks directly from the asset pricing equation (4.2) (Claus and Thomas, 2001).[28] This model is, however, not consistent with our framework as it

[27] Our results are robust to lowering the forecast horizon (and thus the implied average stock-holding period) to five years.

[28] With a measure of the expectations about future real cash flows and under the additional assumption that the time t required return remains constant in the future, one can infer $E_t[R_{t+1}] = E_t[R_{t+2}] = \ldots$ from (4.2). Along those lines, we obtain annual estimates of the DDM implied equity premium from Damodaran (2015). These estimates are then transformed to the quarterly frequency using the BAA-AAA default spread. Specifically, we compute the

Figure 4.4.1: Estimated expected 10-year equity premium and excess dividends



Notes: The figure shows the estimated expected 10-year equity premium (left axis, in percent) and the estimated expected excess dividends (right axis) obtained from forecasts of the auxiliary VAR.

excludes the possibility of asset mispricing and assumes that the transversality condition holds at all times. Therefore, we only employ it as a robustness check of our proxy of the expected equity premium, and presume a positive correlation between the two measures in case mispricing is only weak or restricted to some periods.

We find that our forecast measure of the expected equity premium derived from the auxiliary VAR model is strongly correlated with both the BAA-AAA spread (correlation of 0.64) and the measure derived from the DDM (correlation of 0.49). Figure 4.C.5 plots the three (standardized) measures for expected risk premia. The figure shows that the trend in all proxies is similar, with rising expected premia until about 1985, a following downward trend, and a strong spike around the global fi-

ratio of the DDM equity premium and the annual average of the default spread. Assuming that this ratio remains constant over one year, the spread-adjusted DDM equity premium can then be obtained as the average quarterly BAA-AAA spread multiplied by that ratio.

nancial crisis. However, the decrease of the expected equity premium obtained from our auxiliary VAR is more pronounced during the dot-com bubble, and the rise during the financial crisis is less strong compared to the other two measures. Further, our preferred expectations measure appears more stable overall. The frequent sharp rises and falls in the DDM premium may reflect the fact that it is potentially distorted by the presence of stock mispricing.

Given our measure for the rational expectations on excess dividends, the equity premium and the observations for total stock prices, we can evaluate the implied mispricing component contained in the S&P 500. For this, we first discount at each margin $t = 1, ..., T - h$ the next h realizations of excess dividends and our measure for the expected equity premium. From equation (4.7), we can then obtain $\psi_t = p_t - \left(E_t\left[\sum_{i=1}^{\infty} \rho^{i-1} d^e_{t+i}\right] - E_t\left[\sum_{i=1}^{\infty} \rho^{i-1} r^e_{t+i}\right]\right)$, where we approximate the infinite sum by the next $h = 20$ realizations.[29] The approximated mispricing component is displayed alongside the log real S&P 500 and the implied risk-adjusted fundamental component (the sum of future discounted excess dividends less the sum of future discounted expected equity premia) in Figure 4.C.6.

Importantly, the level of mispricing is indeterminate due to the omitted constant in equation (4.7). Nonetheless, the figure allows us to evaluate the evolution of the mispricing component relative to the dynamics of the log real S&P 500 index. Thus, the figure reveals some time variation in the contributions of the risk-adjusted fundamental component and the mispricing component to the overall stock price. Up to 1982, the size of the risk-adjusted fundamental component exhibits a steady downward trend while stock prices fluctuated more strongly, first showing an upward trend until 1975 and then starting to decrease thereafter. Thus, most of the short-run dynamics of the stock price index can be attributed to a volatile mispricing component. From 1982 onwards, both the fundamental and the mispricing component increase and contributed to the pronounced increase in the overall stock index. Yet, around 1986 – prior to the 1987 stock market crash – the trend growth in the S&P 500 decouples from the relatively stable evolution of the fundamental component and can largely be attributed to a strong increase in the mispricing component. This is again the case during the run-up to and the crash of the dot-com bubble in the second half of the 1990s and the early 2000s. Thus, our measure of

[29] The implied mispricing is robust to extending the approximation of the infinite sum to $h > 20$ observations.

the mispricing component coincides well with anecdotal evidence of excessive stock price periods.

In the following, we refrain from using the estimated mispricing component directly in our VAR model, but infer the impulse response of the mispricing component to a monetary policy shock from the responses of the observed stock price index, and of the expectations about future excess dividends and excess equity returns. This allows a comparison of our results with those of Galí and Gambetti (2015). However, we find that our results are robust to including the estimated mispricing component in the VAR instead of the observed stock price index and our expectation measures, and estimating the mispricing response directly.

4.5 Results

We first present the estimated responses of the stock price and its components to a contractionary monetary policy shock in the framework of a constant coefficient VAR (C-VAR), which allows to assess the significance of the findings more efficiently than the TVC-VAR.[30] Later, we move to the TVC-VAR results and show that the TVC-VAR supports the general findings from the C-VAR. However, there is some evidence for time-variation in the responses of stock prices and expected equity premia to monetary shocks. The monetary policy shock is calibrated to raise the policy rate by 100 basis points (BPS) at each point in time to isolate changes in the transmission from changes in the size of the shock for the TVC-VAR. The individual stock price components are calculated as suggested by our stock price equation (4.7), where we approximate the infinite sum of discounted excess dividends and expected equity premia by the sum over the next 20 periods.

[30] We estimate both models with Bayesian methods using the Gibbs sampler algorithm. To approximate the posterior distribution of the C-VAR we obtain 10,000 draws (after a burn-in of 5,000 draws) of which we retain every fifth to remove potential serial correlation. For the TVC-VAR, we increase the burn-in to 10,000 draws and approximate the posterior by every fifth of 25,000 draws. In both models, we obtain the admissible set of impulse responses by assessing 10,000 candidate impact response matrices to fulfill the zero and sign restrictions for each draw of the reduced form parameters using the algorithm of Binning (2013).

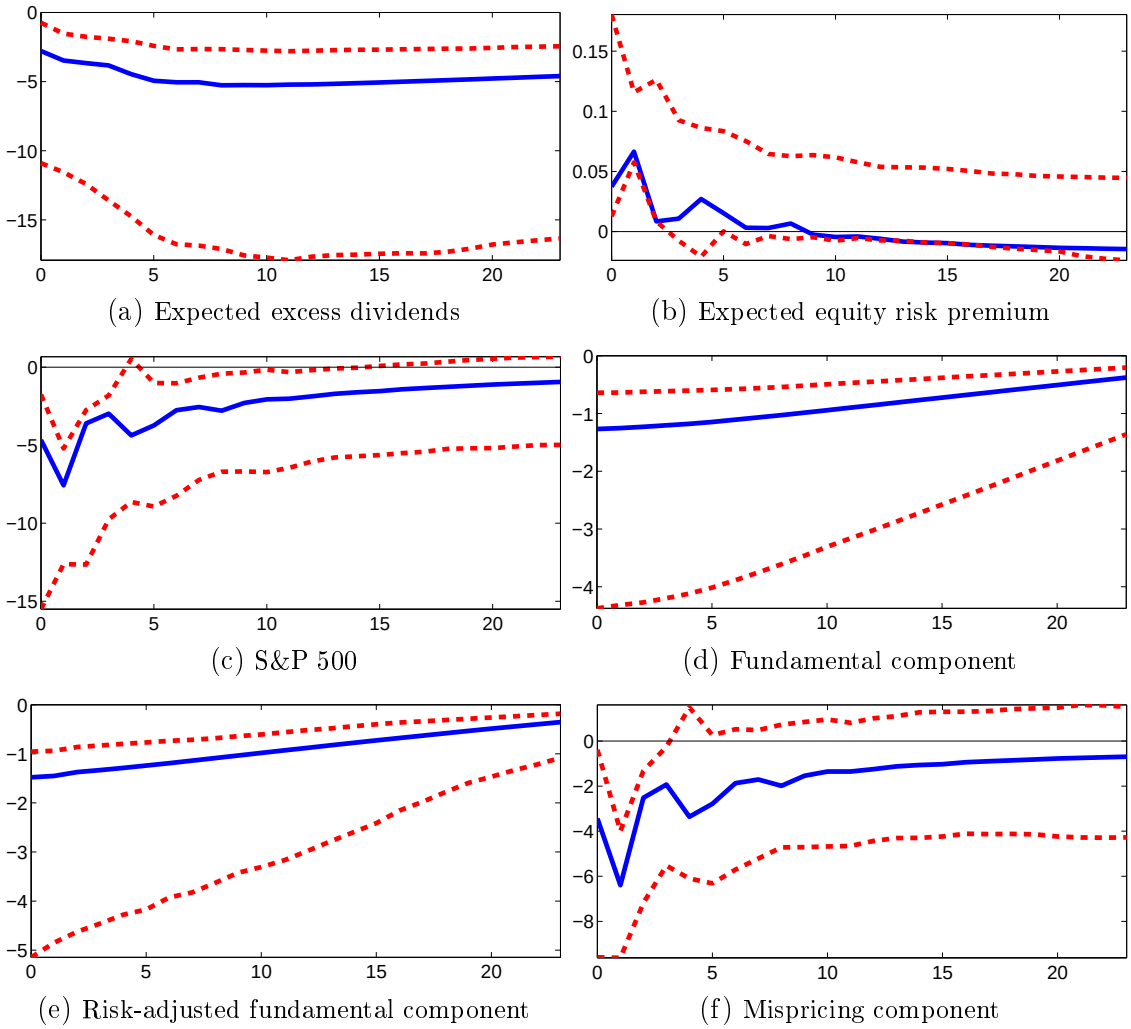
4.5.1 Evidence from a constant coefficient VAR

Figure 4.5.2 shows the cumulative impulse responses of the financial variables and the inferred responses of the individual stock price components. In detail, Panels 4.5.2a-4.5.2c show the responses of the variables that enter our VAR model, i.e. the measures for expected excess dividends and the risk premium, as well as the real S&P 500 price index. To evaluate the contribution of the fundamental components to the total stock price response to a monetary policy shock, we display the response of the three individual stock price components as inferred from equation (4.7) in Panels 4.5.2d-4.5.2f. In particular, we assess the impulse responses of the fundamental component, given by the expected sum of future discounted excess dividends, the fundamental component adjusted for the expected sum of future discounted equity premia (labeled risk-adjusted fundamental component), and, finally, the mispricing component that is obtained by subtracting the response of the risk-adjusted fundamental component from the observed total response of stock prices.

We find that expectations about excess dividends decrease on impact by about 2.9 percent and the expected equity risk premium increases by about 3.7 BPS (Panels 4.5.2a and 4.5.2b). Both responses are significant for at least three quarters. Further, we find a strong negative response of the real S&P 500 index of about five percent on impact which is significant for about three years (Panel 4.5.2c). While our point estimate is smaller in absolute terms than the nine percent estimate of Bjørnland and Leitemo (2009), their estimate lies well within our posterior error bands. It is worth mentioning that when we repeat our estimations relying on the more restrictive Cholesky identification scheme of Galí and Gambetti (2015), we cannot replicate the negative response of the stock market index. Hence, we argue that the identification strategy plays an important role in explaining the differences in the estimation results, where ours is less restrictive. However, also the addition of both expectation measures adds to our finding that stock prices are suppressed for a prolonged period after the shock.

Inferred from the response of expected excess dividends, we, similarly to Galí and Gambetti (2015), find that the fundamental component decreases significantly by about 1.3 percent on impact (Panel 4.5.2d). This decline is persistent and significant throughout. Yet, this response implies that the fundamental factor explains less than half of the immediate decrease in the S&P 500. Subtracting the implied response of

Figure 4.5.2: Impulse responses of the S&P 500 and its components to a monetary policy shock



(a) Expected excess dividends

(b) Expected equity risk premium

(c) S&P 500

(d) Fundamental component

(e) Risk-adjusted fundamental component

(f) Mispricing component

Notes: The figure shows the cumulative responses (in percent) of expected excess dividends, the expected equity premium, the S&P 500 price and its implied components to a 100 BPS contractionary monetary policy shock for impulse response horizons of up to 24 months (x-axis). The solid blue lines indicate the impulse response functions of the median-target model obtained by minimizing the absolute deviation to the pointwise median responses of all variables to the monetary policy shock for all impulse horizons. The dashed red lines denote the 68% posterior error bands following Kilian and Murphy (2014).

the required risk premium, we find that the explanatory content increases further and that the responses in excess dividends and the expected risk premium together account for a 1.5 percent fall in stock prices (Panel 4.5.2e).

Thus, this leaves around 3.5 percentage points of the observed five percent impact drop in the S&P 500 index unexplained. From equation (4.7) this unexplained part can hence be attributed to the response of the mispricing component as displayed in Panel 4.5.2f. Consequently, we conclude that a contractionary monetary policy shock systematically lowers stock prices over and beyond what is implied by fundamental factors. We find this overreaction to be significant for about three quarters.

This implies that a central bank could use such a policy to lean against the wind of stock price bubbles if it perceives stock prices to be excessively high.[31] Further, our result contrasts the findings of Galí and Gambetti (2015) and is difficult to align with the theoretical predictions of the rational bubble framework. In contrast, these findings can be explained by a framework in which stock mispricing arises from false subjective beliefs of irrational investors.

We assess the robustness of our results along two dimensions. First, our results are robust to relaxing the sign restriction on expected excess dividends. When we allow for a positive response, expected excess dividends do not decrease significantly across the studied impact horizons. Since the response of all other financial variables is largely unaffected, this implies a smaller response of the fundamental component and thus an even larger negative response of the mispricing component to a contractionary monetary policy shock. Moreover, we find our results to be robust to employing the BAA-AAA spread or the DDM implied equity return as alternative measures for expected excess returns. These results are displayed in Figures 4.C.7 and 4.C.8. We find a larger response of the approximated expected equity premium of twelve BPS for the BAA-AAA spread (Panel 4.C.7b) and 17 BPS for the DDM-implied premium (Panel 4.C.8b) compared to the VAR-forecast measure. Furthermore, the total response of the S&P 500 is somewhat smaller at around 4.5 percent when using the BAA-AAA spread. For this specification, we thus find the explanatory power of the mispricing component to be smaller and only significant after impact (Panel 4.C.7f). In contrast, when we employ the DDM implied equity return as our measure for expected equity risk premia the results are quantitatively similar to our benchmark estimation, yet the response of the mispricing component is significant for three years (Panel 4.C.8f).

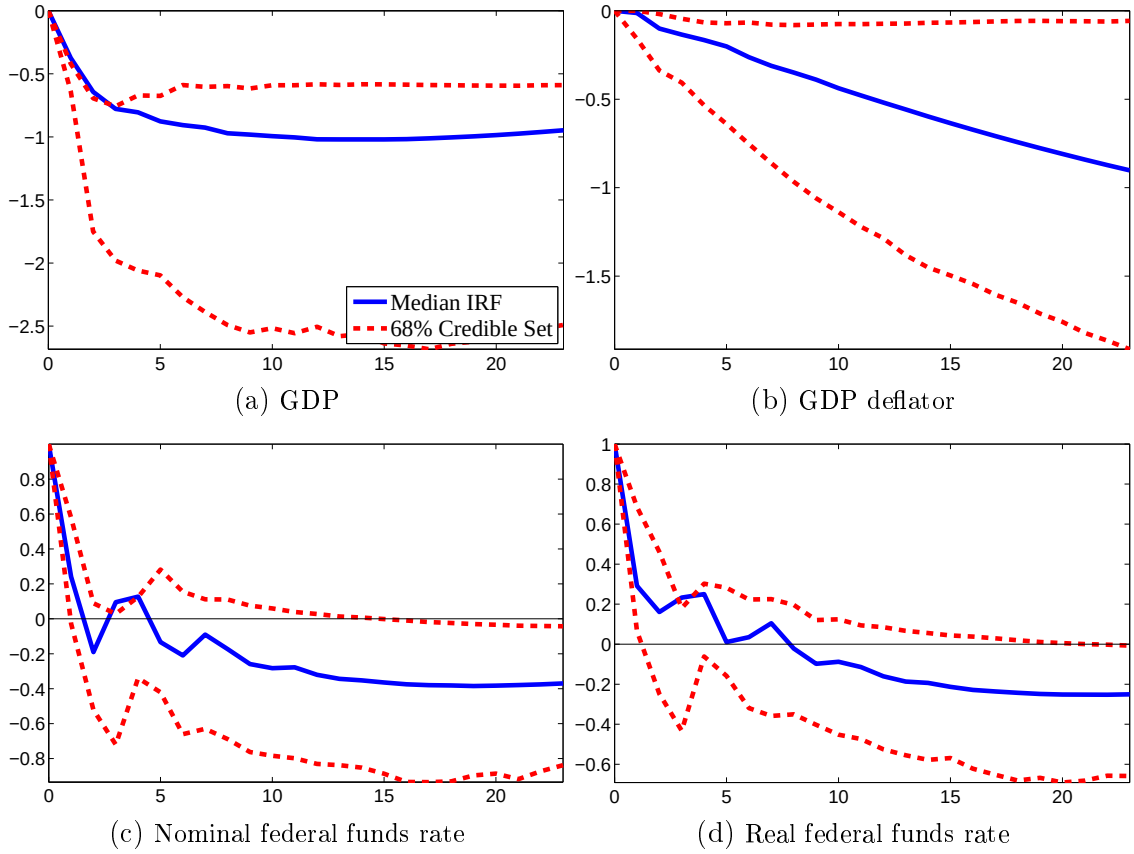
[31] It is important to note that our results are subject to the Lucas' critique. If monetary policy were to exploit these findings in a systematical manner, investors will take this into account, which may already reduce the frequency of excessive stock price periods.

Our findings suggest that central banks *could* attenuate excessive stock price developments by raising the policy rate. Yet, in order to assess whether central banks also *should* make use of such a LATW policy, we need to trade off the benefits of such a policy against its real economic costs. Figure 4.5.3 shows the cumulative responses of output and inflation to the exogenous 100 BPS increase in the policy rate (Panels 4.5.3a-4.5.3b). The figure reveals that a leaning against the wind policy lowers output significantly by one percent after two years. Similarly, prices decrease significantly by about one percent in the long run. Deflating excessively high stock prices through conventional interest rate policy puts considerable downward pressures on real economic activity. In times when stock mispricing is *not* accompanied by a positive output gap and inflation above target, this would indicate a conflict between financial stabilization on the one hand, and price and output stabilization on the other hand. Thus, our results feed into the debate of whether monetary policy is too blunt an instrument to stabilize asset prices and, if not, what other policy instruments, e.g. macroprudential policies, are more suitable for this objective. However, one also has to consider that the immediate costs of lower output and inflation could be offset if the relevant counterfactual is that an emerging bubble is not deflated through active monetary policy, but results in a financial market crash with a potentially larger output and inflation contraction. This, however, cannot be captured and estimated in our modeling framework.[32]

Finally, we also address the question how the U.S. FED responded to stock price shocks, which is key for the identifying restriction in Galí and Gambetti (2015). These results are presented in Figure 4.C.9. As can be seen from Panel 4.C.9c, we cannot reject the assumption of Galí and Gambetti (2015) that the FED has not responded to stock market shocks on impact. However, the 68% confidence level bands indicate the considerable uncertainty around this estimate with a larger probability being assigned to an active, countercyclical response of monetary policy to stock price shocks. Overall, these results question the validity of the imposed zero restriction of Galí and Gambetti (2015).

[32]For a concluding statement on the cost-benefits-analysis of such a policy, one would need to consider a general equilibrium model with a loss function of the monetary authority, which is beyond the scope of this paper.

Figure 4.5.3: Impulse responses of real economic variables to a monetary policy shock



(a) GDP (b) GDP deflator

(c) Nominal federal funds rate (d) Real federal funds rate

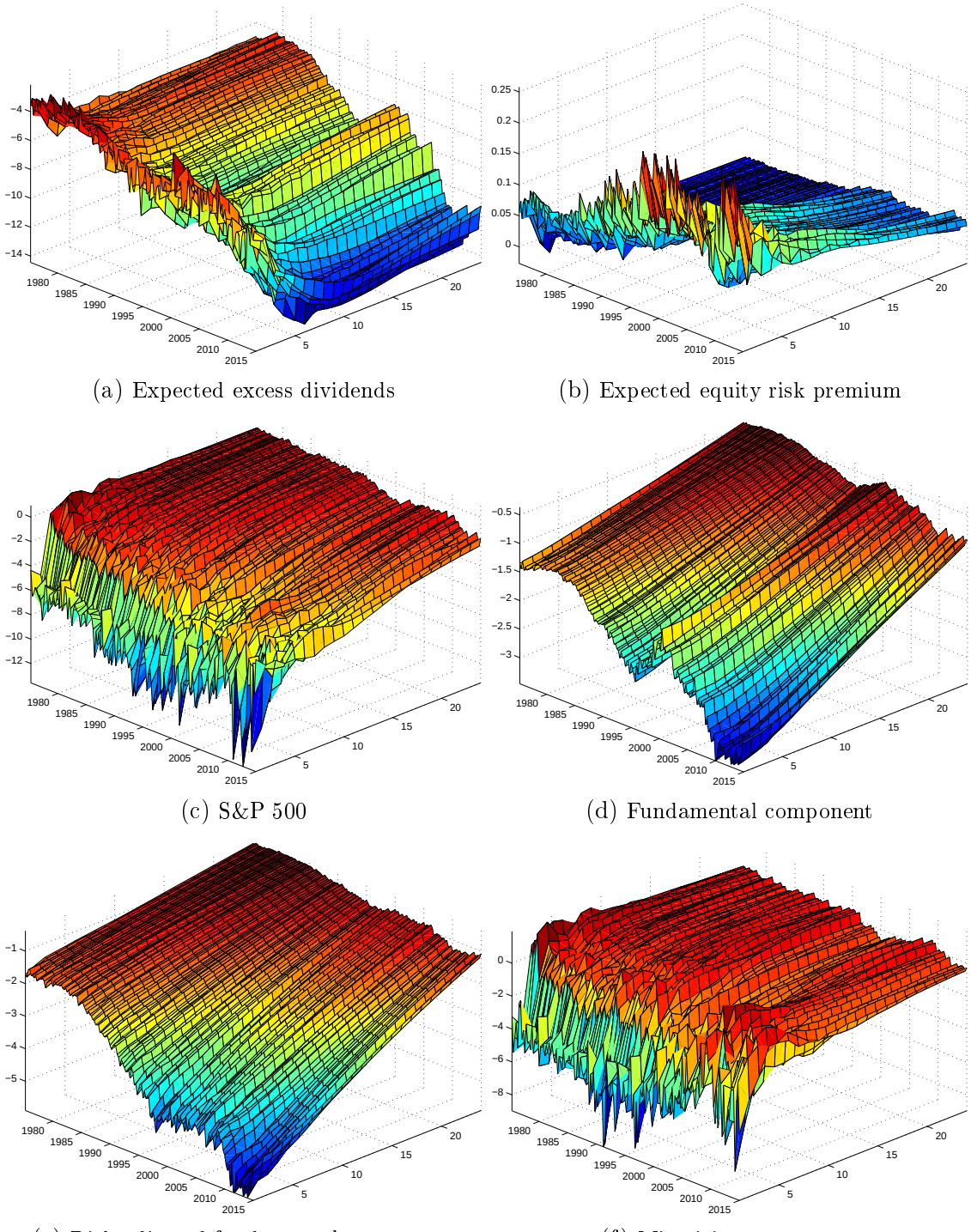
Notes: The figure shows the cumulative responses (in percent) of output, inflation, and the nominal and real policy rates to a 100 BPS contractionary monetary policy shock. See Figure 4.5.2 for further notes.

4.5.2 Evidence from the TVC-VAR

Figure 4.5.4 shows the responses of expected excess dividends, the equity premium, stock prices and its decompositions to a 100 BPS contractionary monetary policy shock over time. In general, the TVC-VAR confirms the main findings from our C-VAR model, but we nevertheless find some time-variation in the impulse responses, which helps shed light on the question of whether a LATW policy is particularly effective during specific periods of time, especially in times of stronger mispricing.

Panel 4.5.4a shows that the negative response of expected excess dividends was quite stable over time and increased in absolute terms only towards the end of the sample. For the expected equity risk premium, time variation in the response is

Figure 4.5.4: Time-varying impulse responses to a monetary policy shock



(a) Expected excess dividends

(b) Expected equity risk premium

(c) S&P 500

(d) Fundamental component

(e) Risk-adjusted fundamental component

(f) Mispricing component

Notes: The figure shows the cumulative, time-varying responses (in percent) of expected excess dividends, the expected equity risk premium, and the S&P500 index, as well as its implied components, to a 100 BPS increase in the federal funds rate. The impulse response functions for up to 24 months (right axis) are obtained from the median-target model.

more pronounced. Panel 4.5.4b shows that until around 1990, the positive response of the expected equity premium was very modest at around five BPS. Thereafter, the response increased considerably up to 20 BPS and became more volatile. Similarly, the response of the stock price index shown in Panel 4.5.4c shows a substantial degree of time variation around 1990. Until then, the S&P 500 decreased by around five percent following a monetary tightening of 100 BPS. Thereafter, however, the response increased in absolute value to around twelve percent during the dot-com bubble in the late 1990's and the last quarters in 2014.

Correspondingly, the individual components also respond more strongly during the latter part of the sample. As shown in Panel 4.5.4e, the impact of the implied risk-adjusted fundamental component continuously strengthened, thus increasing its explanatory power for the decline in stock prices from around two percent in 1980 to around six percent at the end of the sample. Together with the overall stock price response in Panel 4.5.4c, this implies a time-varying response of the mispricing component only since about 1990. As displayed in Panel 4.5.4f, we first confirm the negative (median) response of the mispricing component to a contractionary monetary policy shock over the whole sample, supporting the evidence that the central bank could lower excessive price movements in stocks by raising its policy rate. Yet, we observe that the response has increased sharply in absolute value around 1990 and is particularly strong between 1995 and 2001, the period that is associated with the stock price boom during the dot-com bubble. Our evidence thus suggests that a tighter monetary policy stance during this period could have reduced mispricing by about seven percent.[33]

To summarize, the results of the TVC-VAR offer support to the proponents of a LATW policy and indicate that a monetary tightening lowers excessive stock prices. To give an indication of the significance of this finding, Figure 4.C.10 displays the probability of a negative cumulative response of the mispricing component over time at selected horizons from the credible set of accepted impulse responses. For horizons

[33]In contrast to the time-variation in responses of financial variables, the response of output and inflation to monetary policy shocks is relatively stable across time and corresponds to the responses estimated using the C-VAR. We further note that while the median impulse responses from the TVC-VAR are frequently found to be significantly different from zero as in the C-VAR, the observed time-variation in median IRFs is generally not significant. To be specific, the median IRF at time t and horizon h generally lies within the 68% confidence bands of the same variable at the same horizon for any other time t'.

up to one year, this probability is larger than 50 percent and takes values of up to 90 percent for the impact and one quarter horizons and the time period of up to 1990. Interestingly, however, during the 1990 to 2001 period, when the S&P 500 was most likely overpriced, a monetary policy contraction would likely have been more effective at attenuating mispricing for horizons of one year or longer, compared to the earlier part of the sample. On impact, the chance of reducing mispricing was larger at the beginning of the sample than toward the end of the sample.

4.5.3 Robustness analysis: Monetary policy at the zero lower bound

In response to the global financial crisis, the Federal Reserve Bank gradually lowered its target range for the federal funds rate until it reached the zero-lower bound (ZLB) in December 2008. From this time onwards, the federal funds rate could no longer be used as an instrument to further stimulate the economy. To evaluate whether the estimated time-variation in the effects of monetary policy over time was subject to this constraint, we perform two robustness checks. First, and similar to Galí and Gambetti (2015), we have repeated all estimations on a shorter sample up to 2007Q4, thus excluding the global financial crisis, and find that our results are largely robust to the shorter sample period. Second, we run our estimations using the Wu and Xia (2014) shadow rate instead of the federal funds rate in our VAR. The shadow rate accounts for unconventional monetary policy effects on the term structure of interest rates in a ZLB environment and is used as a measure of the federal funds rate that would prevail in the absence of the ZLB.

Figure 4.C.11 show the time-varying impulse responses of expected excess dividends, the expected equity premium, stock prices, the risk-adjusted fundamental and the speculative component to a 100 BPS monetary policy shock as measured by an increase in the Wu and Xia (2014) shadow rate. We can generally confirm the previous results obtained for the expected excess dividends and the equity risk premium. However, the impact responses are slightly smaller than in the benchmark model. Moreover, we find that the time-variation in the response of stock prices is now less pronounced (Panel 4.C.11c). In combination, this implies a stronger negative response of the mispricing component in the first part of the sample up until about 1995, and a weaker but still negative response toward the end of the sample. In particular, the dot-com bubble episode no longer stands out. Thus, we

can conclude that the policy constraint at the zero lower bound explains part of the observed time variation of the response of the mispricing component of stock prices.

4.6 Conclusion

In this paper, we explore the effects of monetary policy on stock prices and on their underlying components. By allowing for the possibility of stock mispricing, we address the question whether central banks could implement a leaning against the wind policy to attenuate excessive stock price developments. We find that stock prices decrease significantly and persistently in response to a monetary policy tightening. Decomposing this response into a fundamental component related to rational expectations about future discounted dividends, and an expected equity risk premium, we find that these sources can only account for about one third of the fall in stock prices. Hence, we conclude that the decrease in stock prices can largely be attributed to a systematic negative response of the mispricing component. By this, we provide support to the claims of the proponents of an active, leaning against the wind monetary policy. If stocks are overpriced, contractionary monetary policy could be used to lower mispricing. Yet, this comes at the cost of considerable downward pressure on prices and real economic activity.

In contrast to previous literature, we therefore argue that such a policy can be implemented without trade-off if output and inflation are also above the central bank's targets. Yet, when excessive stock mispricing arises at times when output and inflation are below their targets, it is debatable whether conventional interest rate policy are not too blunt a tool to dampen excessive stock mispricing in a cost-effective manner. In particular, at these times, the central bank faces the trade-off between containing an excessive asset price bubble and stimulating real economic activity. Given the relatively modest response of the mispricing component to a large monetary policy shock, it is questionable whether conventional monetary policy is the right tool to use. Nonetheless, since we find a significant overreaction of stock markets to interest rate shocks, policy makers need to carefully consider financial stability concerns also for accommodative monetary policy measures, as they may induce a significant and persistent overpricing in stock markets, and foster the build-up of imbalances in the financial system.

Finally, our results are difficult to align with a concept of mispricing that assumes that all agents are rational and hold objective expectations about future dividends and excess capital gains. Such a framework employed by previous literature predicts that the mispricing component increases in response to a contractionary monetary policy shock in the long-run. In contrast, our findings are in line with a framework in which a share of investors are irrational and hold false subjective expectations about the stock's underlying fundamentals. Such a framework may arise from a range of possible behavioral explanations.

Appendix

4.A Log-linear approximation of asset pricing equation

Under the assumption that the price-dividend ratio is stationary, the log-linear approximation of equation (4.1) is given by:

$$\begin{aligned} r_{t+1} &= \log(P_{t+1} + D_{t+1}) - \log(P_t) \\ &= p_{t+1} - p_t + \log\left(1 + \exp(d_{t+1} - p_{t+1})\right), \end{aligned} \tag{4.17}$$

where logs of variables are denoted by lowercase letters and $r_{t+1} = \log(1 + R_{t+1})$. Log-linearizing equation (4.17) around the steady state using a first-order Taylor expansion, we obtain:

$$r_{t+1} = k + \rho p_{t+1} + (1 - \rho) d_{t+1} - p_t, \tag{4.18}$$

where ρ and k are parameters of linearization defined as $\rho \equiv 1/\left(1 + \exp(\overline{d - p})\right)$, where $\left(\overline{d - p}\right)$ is the average log dividend-price ratio (such that $0 < \rho < 1$), and $k \equiv -\log(\rho) - (1 - \rho)\log(1/\rho - 1)$.

Solving equation (4.18) for p_t and iterating forward, the stock price can be written as a linear combination of future dividends, future stock returns, and a terminal value (disregarding the constant term k):

$$p_t = \sum_{i=1}^{\infty} \rho^{i-1} \left[(1 - \rho) d_{t+i} - r_{t+i}\right] + \lim_{T \to \infty} \rho^T p_{t+T}. \tag{4.19}$$

4.B Priors and Estimation of TVC-VAR

We follow Primiceri (2005) and Galí and Gambetti (2015) and assume that the initial states for the coefficients θ_t, the covariances α_t, the log volatilities σ_t and the hyperparameters Q, V and W are independent of each other. The priors for the initial states of the coefficients, the covariances and the log standard errors are assumed to be normally distributed. The priors of the hyperparameters are assumed to be distributed as independent inverse-Wishart. All these assumptions are standard in the literature as described in Primiceri (2005). The priors thus take

the specific forms:

$$\begin{aligned}
\theta_0 &\sim N(\hat{\theta}, 4 * \widehat{Cov}(\theta)) \\
\alpha_{i0} &\sim N(\hat{\alpha}, \widehat{Cov}(\alpha_i)) \\
\log(\sigma_0) &\sim N(\log(\hat{\sigma}), I_K) \\
Q^{-1} &\sim IW(k_Q(K+1)\widehat{Cov}(\theta), K+1) \\
V_i^{-1} &\sim IW(k_V(i+1)\widehat{Cov}(\alpha_i), i+1) \\
W^{-1} &\sim IW(k_W(K+1)I_K, K+1)
\end{aligned}$$

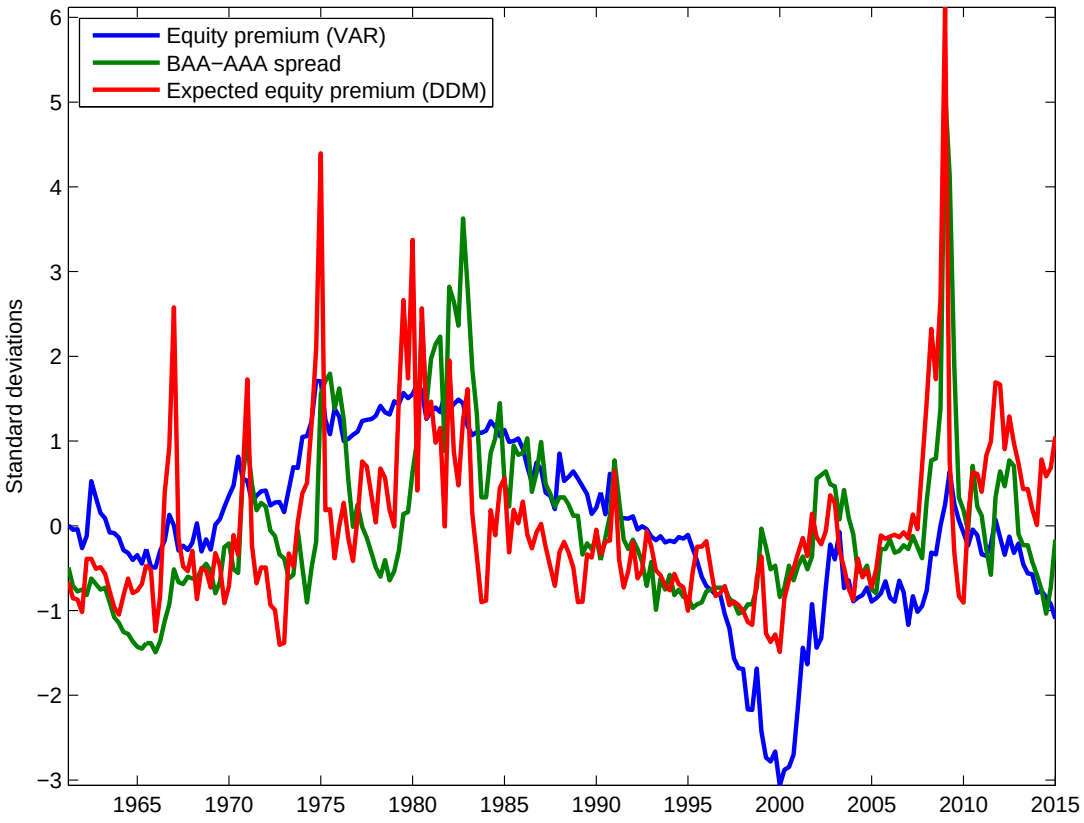
where the prior means and covariances are obtained from OLS-estimates from a constant VAR for a training sample with $T_0 = 48$ observations and with $k_Q = 0.005$, $k_V = 0.01$, $k_W = 0.01$ so that the priors are not flat but diffuse. Note that the prior V is an inverse-Wishart with degrees of freedom and scale parameter adjusted to the number of variables in the respective equation $i = 2, \ldots, n$. For the constant VAR specification, we impose independent Normal-Wishart priors with identical means, variances and degrees of freedom for θ_0 and H_0 as for the TVC-VAR specification.

The model is estimated using the Gibbs sampling algorithm described in Del Negro and Primiceri (2015). A summary is available in the online appendix of Galí and Gambetti (2015).

Appendix

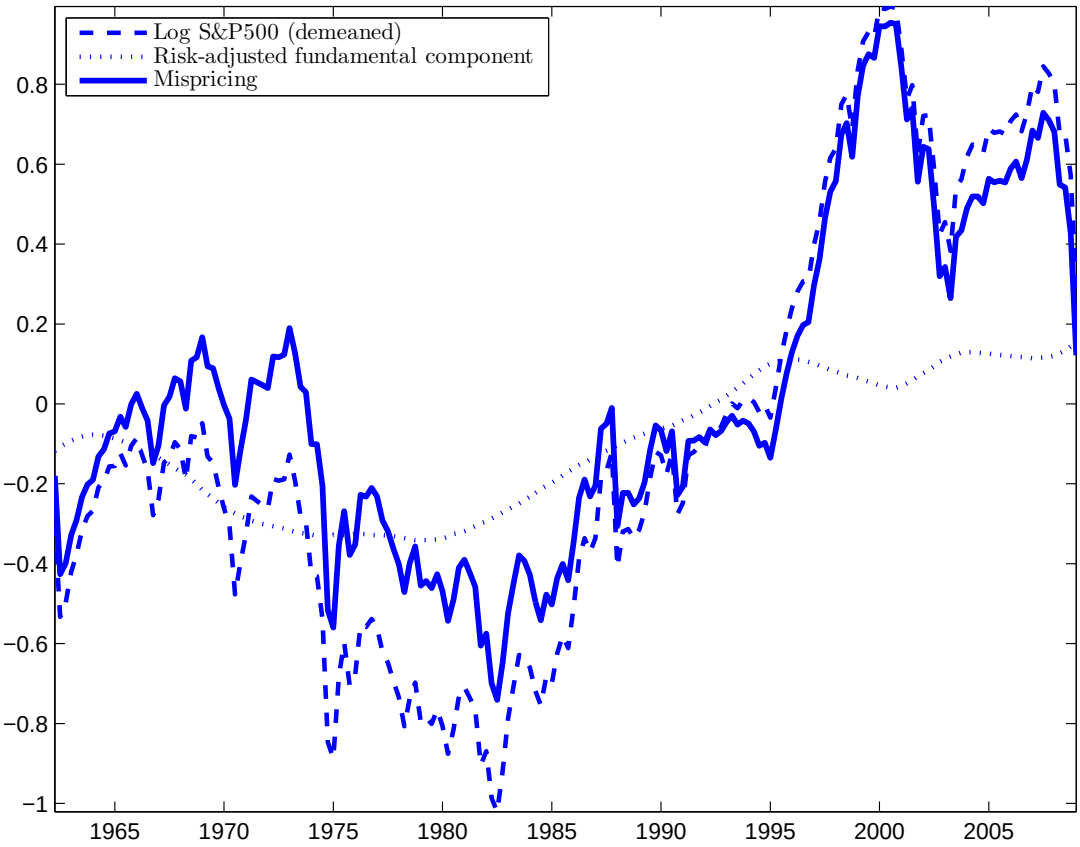
4.C Figures

Figure 4.C.5: Comparison of standardized proxies for the expected 10-year equity premium



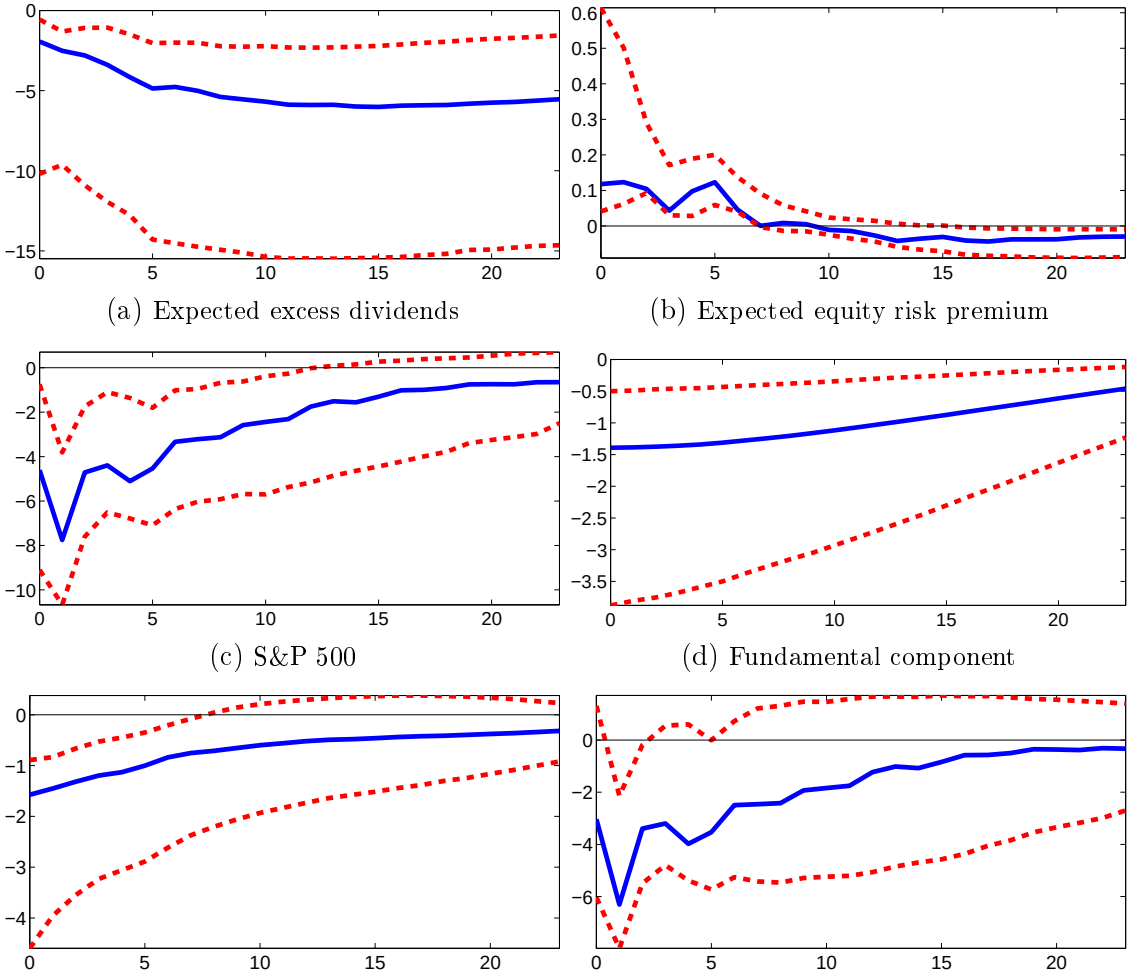
Notes: The figure compares the estimated expected equity premium from the auxiliary VAR (blue line) with other available proxies for the expected 10-year equity premium. These are the BAA-AAA spread (green line) and the expected equity premium from a dividend discount model (red line).

Figure 4.C.6: Implied mispricing in the S&P 500



Notes: Log real S&P500 index (demeaned) and implied risk-adjusted fundamental and mispricing components from equation (4.7).

Figure 4.C.7: Impulse responses of the S&P 500 and its components to a monetary policy shock with the BAA-AAA spread



(a) Expected excess dividends

(b) Expected equity risk premium

(c) S&P 500

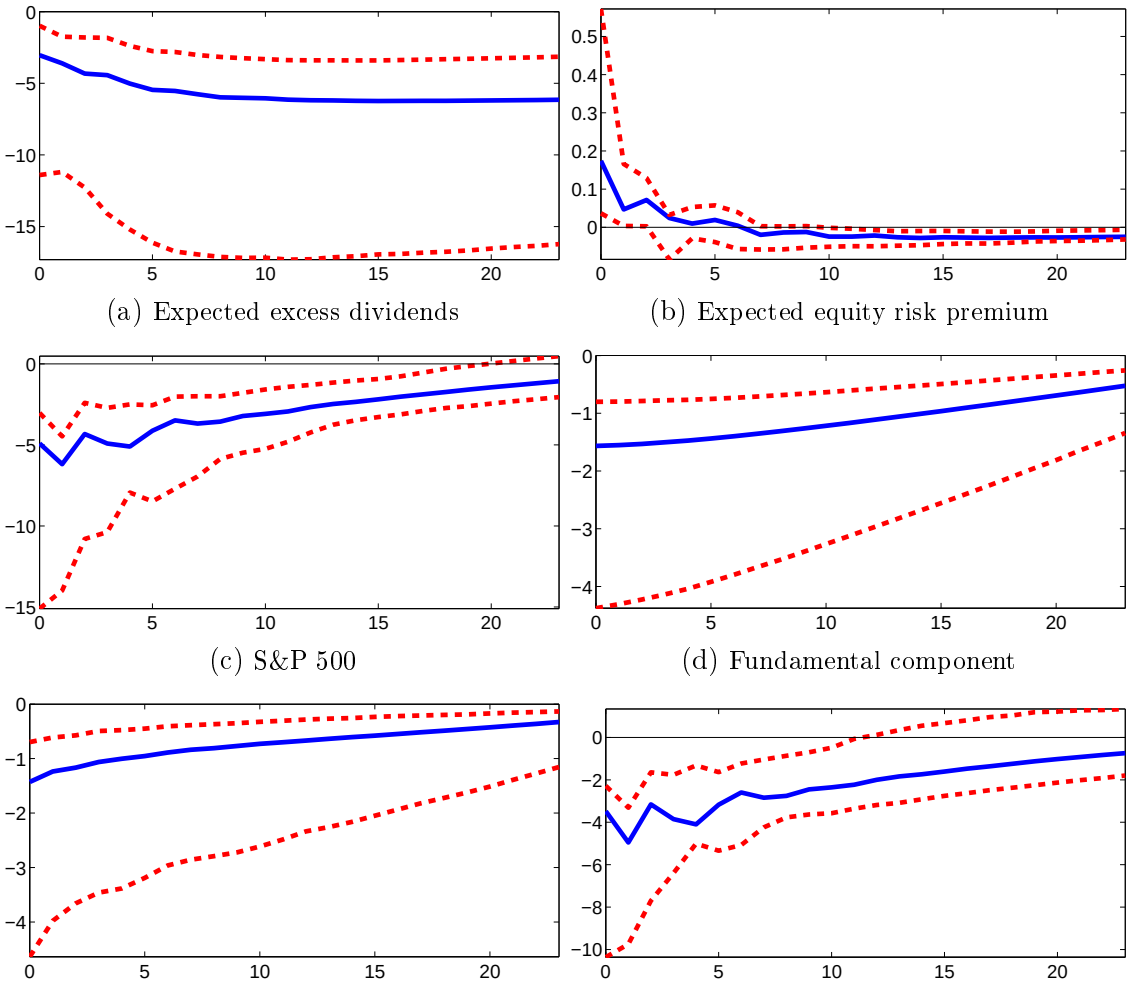
(d) Fundamental component

(e) Risk-adjusted fundamental component

(f) Mispricing component

Notes: The figure shows the cumulative responses (in percent) of expected excess dividends, the expected equity premium, the S&P500 price and its implied components to a 100 BPS increase in the federal funds rate, where the expected equity risk premium is measured as the BAA-AAA spread. See Figure 4.5.2 for further notes.

Figure 4.C.8: Impulse responses of the S&P 500 and its components to a monetary policy shock the with DDM equity premium



(a) Expected excess dividends

(b) Expected equity risk premium

(c) S&P 500

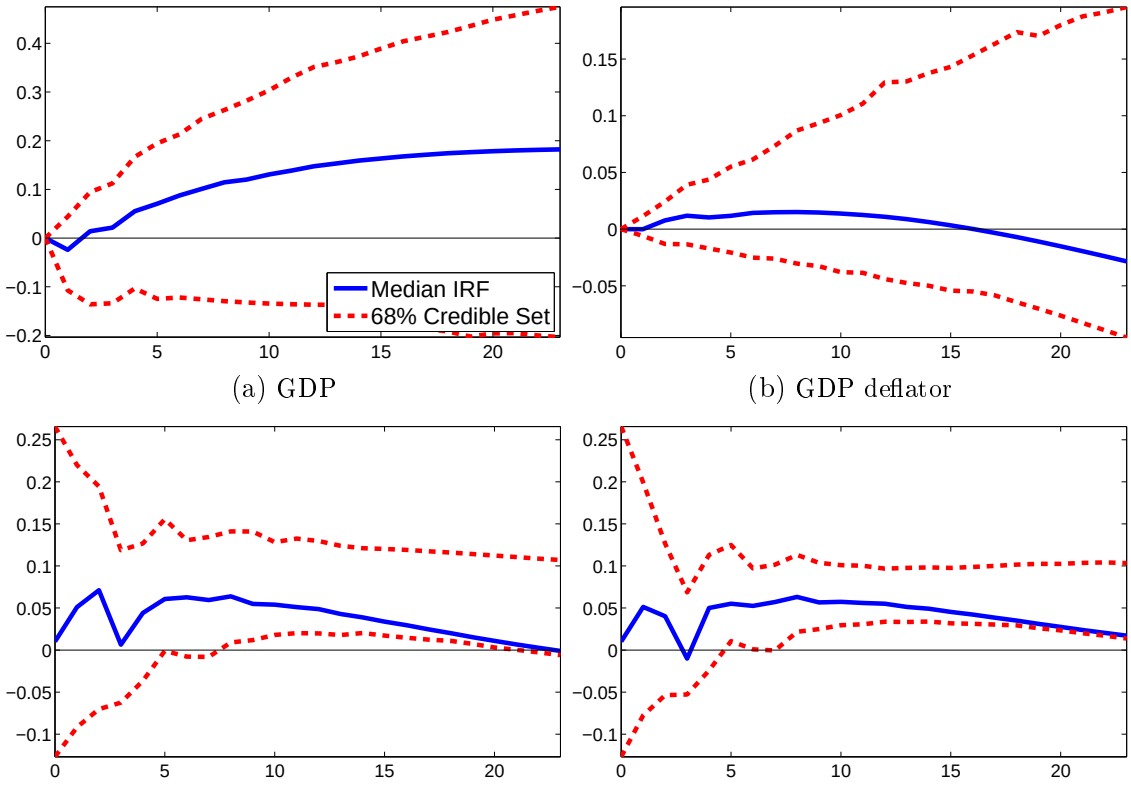
(d) Fundamental component

(e) Risk-adjusted fundamental component

(f) Mispricing component

Notes: The figure shows the cumulative responses (in percent) of expected excess dividends, the expected equity premium, the S&P500 price and its implied components to a 100 BPS increase in the federal funds rate, where the expected equity premium is obtained from the dividend discount model (DDM). See Figure 4.5.2 for further notes.

Figure 4.C.9: Impulse responses of real economic variables to a stock price shock



(a) GDP

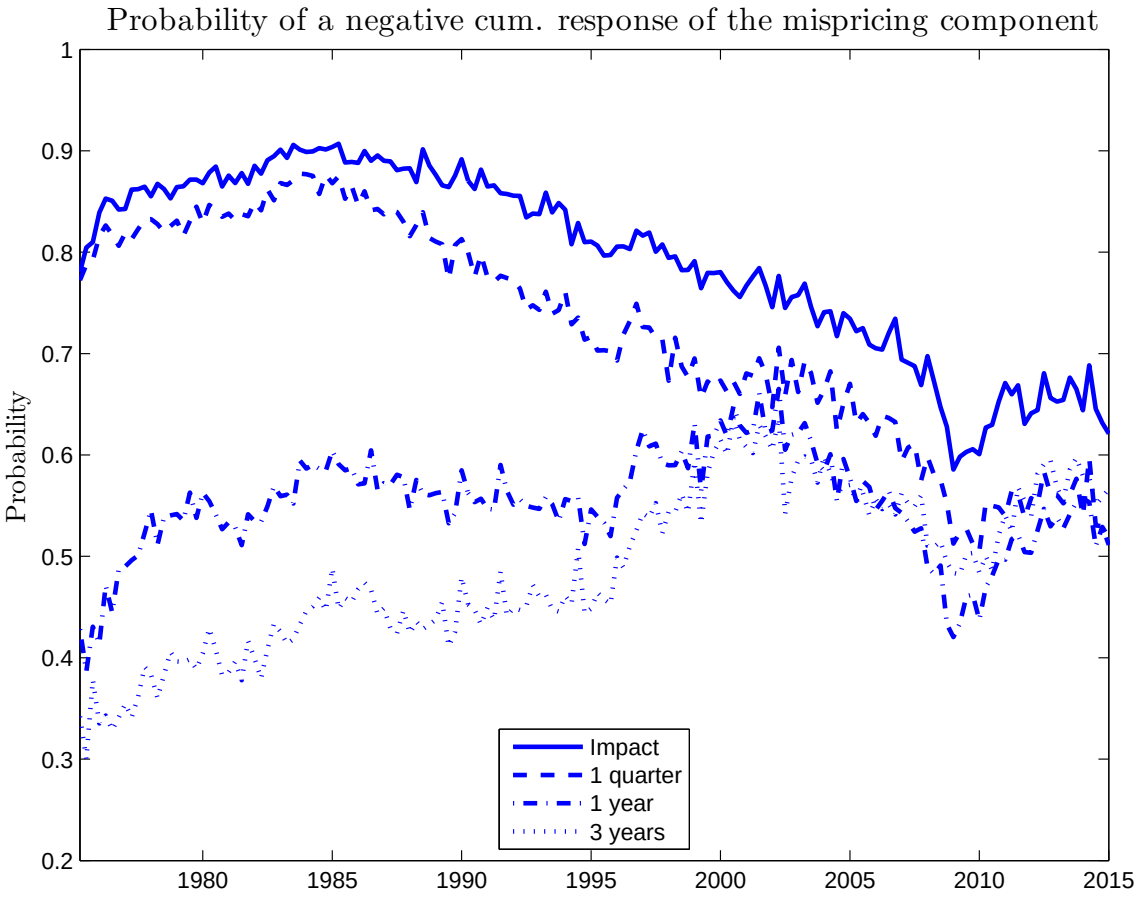
(b) GDP deflator

(c) Nominal federal funds rate

(d) Real federal funds rate

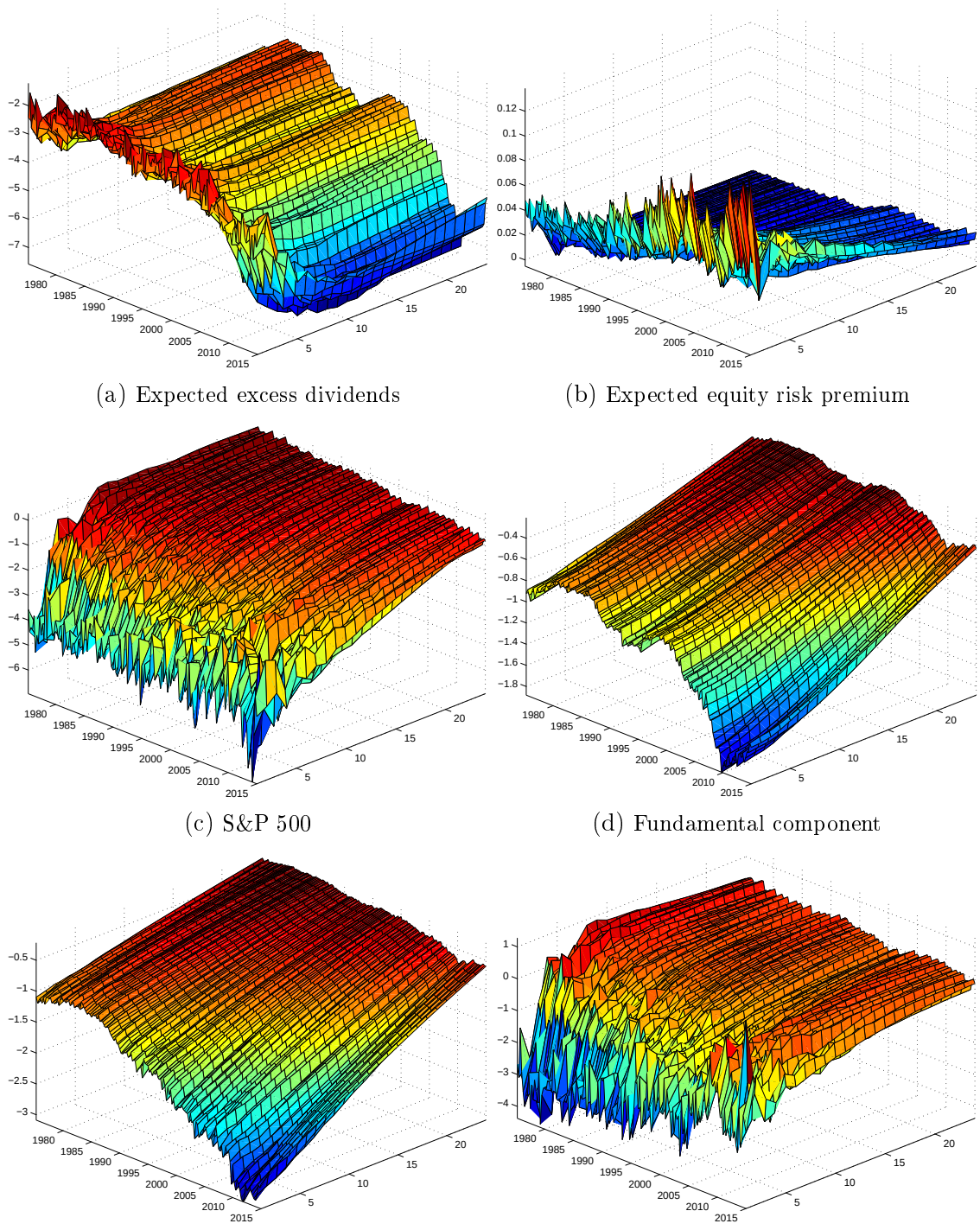
Notes: The figure shows the cumulative responses (in percent) of real economic variables and the nominal and real policy rates to an exogenous one percent stock price increase. See Figure 4.5.2 for further notes.

Figure 4.C.10: Time-varying probability of a negative mispricing response to a monetary policy shock



Notes: The figure shows the probability of a negative cumulative response of the mispricing component to a 100 BPS increase in the federal funds rate over time at selected horizons.

Figure 4.C.11: Time-varying impulse responses to a Wu and Xia (2014) shadow rate monetary policy shock



(a) Expected excess dividends

(b) Expected equity risk premium

(c) S&P 500

(d) Fundamental component

(e) Risk-adjusted fundamental component

(f) Mispricing component

Notes: The figure shows the cumulative, time-varying responses (in percent) of expected excess dividends, the expected equity risk premium, and the S&P500 index, as well as its implied components, to a 100 BPS increase in the Wu and Xia (2014) shadow rate. The impulse response functions are obtained from the median-target model.

Bibliography

Adalid, R. and C. Detken (2007): Liquidity Shocks and Asset Price Boom/Bust Cycles, Working Paper No. 732, European Central Bank.

Adam, K., A. Marcet and J. Beutel (2015): Stock Price Booms and Expected Capital Gains, Working Papers No. 757, Barcelona Graduate School of Economics.

Angelidis, T. and S. Degiannakis (2007): Backtesting VaR Models: A Two-Stage Procedure, *Journal of Risk Model Validation*, 1(2), 27–48.

Arias, J.E., J.F. Rubio-Ramirez and D.F. Waggoner (2014): Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications, Working Paper No. 2014-1, Federal Reserve Bank of Atlanta.

Assenmacher-Wesche, K. and S. Gerlach (2010): Monetary Policy and Financial Imbalances: Facts and Fiction, *Economic Policy*, 25, 437–482.

Bai, J. and S. Ng (2002): Determining the Number of Factors in Approximate Factor Models, *Econometrica*, 70(1), 191–221.

Bai, J. and S. Ng (2008): Forecasting Economic Time Series Using Targeted Predictors, *Journal of Econometrics*, 146(2), 304–317, Honoring the research contributions of Charles R. Nelson.

Basel Committee on Banking Supervision (2005). Amendment to the Capital Accord to Incorporate Market Risk. Bank for International Settlements.

Basel Committee on Banking Supervision (2009). Revisions to the Basel II Market Risk Framework. Bank for International Settlements.

Basel Committee on Banking Supervision (2016). Minimum Capital Requirements for Market Risk. Standards, Bank for International Settlements.

Bates, J.M. and C.W.J. Granger (1969): The Combination of Forecasts, *OR*, 20(4), 451–468.

Baumeister, C. and L. Benati (2013): Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound, *International Journal of Central Banking*, 9(2), 165–212.

Baumeister, C. and G. Peersman (2013): Time-Varying Effects of Oil Supply Shocks on the US Economy, *American Economic Journal: Macroeconomics*, 5(4), 1–28.

Bernanke, B. and M. Gertler (1999): Monetary Policy and Asset Price Volatility, *Economic Review*, Q IV, 17–51.

Bernanke, B.S., M. Gertler and S. Gilchrist (1999): The Financial Accelerator in a Quantitative Business Cycle Framework, *Handbook of Macroeconomics*, 1, 1341–1393.

Bernanke, B.S. and K.N. Kuttner (2005): What Explains the Stock Market's Reaction to Federal Reserve Policy?, *The Journal of Finance*, 60(3), 1221–1257.

Binning, A. (2013): Underidentified SVAR Models: A Framework for Combining Short and Long-Run Restrictions with Sign-Restrictions, Working Paper No. 2013/14, Norges Bank.

Bjørnland, H.C. and K. Leitemo (2009): Identifying the Interdependence between US Monetary Policy and the Stock Market, *Journal of Monetary Economics*, 56(2), 275–282.

Bollerslev, T. (1986): Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31(4), 307–327.

Bollerslev, T. (1990): Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model, *The Review of Economics and Statistics*, 72(3), 498–505.

Bordo, M.D. and J. Landon-Lane (2013): Does Expansionary Monetary Policy Cause Asset Price Booms; Some Historical and Empirical Evidence, Working Paper No. 19585, National Bureau of Economic Research.

Borio, C. and P. Lowe (2002): Asset Prices, Financial and Monetary Stability: Exploring the Nexus, BIS Working Papers No. 114, Bank for International Settlements.

Borio, C. and P. Lowe (2004): Securing Sustainable Price Stability: Should Credit Come Back From the Wilderness?, BIS Working Papers No. 157, Bank for International Settlements.

Broock, W.A., J.A. Scheinkman, W.D. Dechert and B. LeBaron (1996): A Test for Independence Based on the Correlation Dimension, *Econometric Reviews*, 15(3), 197–235.

Brooks, C., S. Burke, S. Heravi and G. Persand (2005): Autoregressive Conditional Curtosis, *Journal of Financial Econometrics*, 3, 399–421.

Brunnermeier, M.K. and C. Julliard (2008): Money Illusion and Housing Frenzies, *Review of Financial Studies*, 21(1), 135–180.

Brunnermeier, M.K. and J.A. Parker (2005): Optimal Expectations, *American Economic Review*, 95(4), 1092–1118.

Bry, G. and C. Boschan (1971): *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs.* NBER Books. National Bureau of Economic Research, Inc.

Cai, Z. (2002): Regression Quantiles for Time Series, *Econometric Theory*, 18, 169–192.

Cai, Z. and X. Wang (2008): Nonparametric Estimation of Conditional VaR and Expected Shortfall, *Journal of Econometrics*, 147(1), 120–130.

Camerer, C. (1989): Bubbles and Fads in Asset Prices, *Journal of Economic Surveys*, 3(1), 3–41.

Campbell, J. and R. Shiller (1988): The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies*, 1(3), 195–228.

Campbell, J.Y. (1991): A Variance Decomposition for Stock Returns, *The Economic Journal*, 101(405), 157–179.

Campbell, J.Y., S. Giglio and C. Polk (2013): Hard Times, *Review of Asset Pricing Studies*, 3(1), 95–132.

Campbell, J.Y., S. Giglio, C. Polk and R. Turley (2012): An Intertemporal CAPM with Stochastic Volatility, NBER Working Papers No. 18411, National Bureau of Economic Research, Inc.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1996): *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

Campbell, J.Y., C. Polk and T. Vuolteenaho (2010): Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns, *Review of Financial Studies*, 23(1), 305–344.

Campbell, J.Y. and R.J. Shiller (1987): Cointegration and Tests of Present Value Models, *Journal of Political Economy*, 95(5), 1062–1088.

Campbell, J.Y. and T. Vuolteenaho (2004): Inflation Illusion and Stock Prices, *American Economic Review*, 94(2), 19–23.

Canova, F. and L. Gambetti (2009): Structural Changes in the US Economy: Is there a Role for Monetary Policy?, *Journal of Economic Dynamics and Control*, 33(2), 477–490.

Cecchetti, S.G., H. Genberg and S. Wadhwani (2002): Asset Prices in a Flexible Inflation Targeting Framework, Working Paper No. 8970, National Bureau of Economic Research.

Chen, X. and Y. Fan (2006): Estimation of Copula-Based Semiparametric Time Series Models, *Journal of Econometrics*, 130(2), 307–335.

Cherubini, U., E. Luciano and W. Vecchiato (2004): *Copula Methods in Finance*, Hoboken N.J.: John Wiley & Sons.

Christiano, L.J., M. Eichenbaum and C.L. Evans (2005): Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, *Journal of Political Economy*, 113(1), 1–45.

Clark, T.E. and M.W. McCracken (2009): Tests of Equal Predictive Ability With Real-Time Data, *Journal of Business & Economic Statistics*, 27(4), 441–454.

Clark, T.E. and K.D. West (2007): Approximately Normal Tests for Equal Predictive Accuracy in Nested Models, *Journal of Econometrics*, 138(1), 291–311.

Claus, J. and J. Thomas (2001): Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets, *The Journal of Finance*, 56(5), 1629–1666.

Cochrane, J. (1992): Explaining the Variance of Price-Dividend Ratios, *Review of Financial Studies*, 5(2), 243–280.

Cochrane, J.H. (2011): How Did Paul Krugman Get It So Wrong, *Economic Affairs*, 31(2), 36–40.

Damodaran, A. (2015): Equity Risk Premiums (ERP): Determinants, Estimation and Implications – The 2015 Edition, , Stern School of Business.

Del Negro, M. and G.E. Primiceri (2015): Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum, *The Review of Economic Studies*, 82(4), 1342.

Detken, C. and F. Smets (2004): Asset Price Booms and Monetary Policy, Working Paper No. 364, European Central Bank.

Diba, B.T. and H.I. Grossman (1988): Explosive Rational Bubbles in Stock Prices?, *The American Economic Review*, 78(3), 520–530.

Diebold, F.X. and R.S. Mariano (1995): Comparing Predictive Accuracy, *Journal of Business & Economic Statistics*, 13(3), 134–144.

Diks, C.G.H., V. Panchenko and D.J.C. van Dijk (2011): Likelihood-Based Scoring Rules for Comparing Density Forecasts in Tails, *Journal of Econometrics*, 163(2), 215–230.

Drost, F. and T. Nijman (1993): Temporal Aggregation of GARCH Processes, *Econometrica*, 61, 909–927.

Engle, R.F. (1982): Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987–1007.

Engle, R.F. and S. Manganelli (2004): CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles, *Journal of Business & Economic Statistics*, 22(4), 367–381.

Escanciano, J.C. (2009): Quasi-maximum Likelihood Estimation of Semi-Strong GARCH Models, *Econometric Theory*, 25, 561–570.

Evans, G.W. (1991): Pitfalls in Testing for Explosive Bubbles in Asset Prices, *American Economic Review*, 81(4), 922–930.

Fisher, R. (1934): *Statistical Methods for Research Workers*, 5^{th} Edition, Edinburgh, U.K.: Oliver and Boyd Ltd.

Flood, R.P. and R.J. Hodrick (1990): On Testing for Speculative Bubbles, *Journal of Economic Perspectives*, 4(2), 85–101.

Friedman, M. (1968): The Role of Monetary Policy, *The American Economic Review*, 58(1), 1–17.

Froot, K.A. and M. Obstfeld (1991): Intrinsic Bubbles: The Case of Stock Prices, *American Economic Review*, 81(5), 1189–1214.

Fry, R. and A. Pagan (2011): Sign Restrictions in Structural Vector Autoregressions: A Critical Review, *Journal of Economic Literature*, 49(4), 938–960.

Furlanetto, F. (2011): Does Monetary Policy React to Asset Prices? Some International Evidence, *International Journal of Central Banking*, 7(3), 91–111.

Galí, J. (2014): Monetary Policy and Rational Asset Price Bubbles, *American Economic Review*, 104(3), 721–752.

Galí, J. and L. Gambetti (2015): The Effects of Monetary Policy on Stock Market Bubbles: Some Evidence, *American Economic Journal: Macroeconomics*, 7(1), 233–257.

Gertler, M. and P. Karadi (2015): Monetary Policy Surprises, Credit Costs, and Economic Activity, *American Economic Journal: Macroeconomics*, 7(1), 44–76.

Glosten, L.R., R. Jagannathan and D.E. Runkle (1993): On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *The Journal of Finance*, 48(5), 1779–1801.

Gordon, M.J. (1962): *The Investment, Financing, and Valuation of the Corporation*, Homewood, Illinois: Richard D Irwin Inc.

Granger, C.W.J. and M.H. Pesaran (2000): Economic and Statistical Measures of Forecast Accuracy, *Journal of Forecasting*, 19(7), 537–560.

Gürkaynak, R.S. (2008): Econometric Tests of Asset Price Bubbles: Taking Stock, *Journal of Economic Surveys*, 22(1), 166–186.

Gust, C. and J.D. López-Salido (2014): Monetary Policy and the Cyclicality of Risk, *Journal of Monetary Economics*, 62, 59–75.

Hamadeh, T. and J.M. Zakoïan (2011): Asymptotic Properties of LS and QML Estimators for a Class of Nonlinear GARCH Processes, *Journal of Statistical Planning and Inference*, 141(1), 488–507.

Hansen, B.E. (1994): Autoregressive Conditional Density Estimation, *International Economic Review*, 35(3), 705–730.

Hansen, P.R., A. Lunde and J.M. Nason (2011): The Model Confidence Set, *Econometrica*, 79(2), 453–497.

Harding, D. and A. Pagan (2002): Dissecting the Cycle: A Methodological Investigation, *Journal of Monetary Economics*, 49(2), 365–381.

Harvey, C.R. and A. Siddique (1999): Autoregressive Conditional Skewness, *Journal of Financial and Quantitative Analysis*, 35, 465–487.

Harvey, D.I., S.J. Leybourne and R. Sollis (2015): Recursive Right-Tailed Unit Root Tests for an Explosive Asset Price Bubble, *Journal of Financial Econometrics*, 13(1), 166–187.

Herwartz, H. (2010): A Note on Model Selection in (Time Series) Regression Models – General-to-Specific or Specific-to-General?, *Applied Economics Letters*, 17(12), 1157–1160.

Herwartz, H. (2013): On the Predictive Content of Autoregression Residuals: A Semiparametric, Copula-Based Approach to Time Series Prediction, *Journal of Forecasting*, 32(4), 353–368.

Himmelberg, C., C. Mayer and T. Sinai (2005): Assessing High House Prices: Bubbles, Fundamentals and Misperceptions, *Journal of Economic Perspectives*, 19(4), 67–92.

Holmstrom, B. and J. Tirole (1997): Financial Intermediation, Loanable Funds, and the Real Sector, *The Quarterly Journal of Economics*, 112(3), 663–691.

Homm, U. and J. Breitung (2012): Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods, *Journal of Financial Econometrics*, 10(1), 198–231.

Kilian, L. and D.P. Murphy (2012): Why Agnostic Sign Restrictions Are Not Enough: Understanding The Dynamics Of Oil Market VAR Models, *Journal of the European Economic Association*, 10(5), 1166–1188.

Kilian, L. and D.P. Murphy (2014): The Role of Inventories and Speculative Trading in the Global Market for Crude Oil, *Journal of Applied Econometrics*, 29(3), 454–478.

Kiyotaki, N. and J. Moore (1997): Credit Cycles, *Journal of Political Economy*, 105(2), 211–248.

Kohn, D.L. (2006). Monetary Policy and Asset Prices. Speech at "Monetary Policy: A Journey from Theory to Practice," a European Central Bank Colloquium held in honor of Otmar Issing, Frankfurt, Germany.

Kupiec, P.H. (1995): Techniques for Verifying the Accuracy of Risk Measurement Models, *Journal of Derivatives*, 3, 73–84.

Kwiatkowski, D., P.C. Phillips, P. Schmidt and Y. Shin (1992): Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, *Journal of Econometrics*, 54(1), 159–178.

Lee, S.W. and B.E. Hansen (1994): Asymptotic Theory for the GARCH(1,1) Quasi Maximum Likelihood Estimator, *Economic Theory*, 10, 29–52.

Lee, T.H. and X. Long (2009): Copula-Based Multivariate GARCH Model with Uncorrelated Dependent Errors, *Journal of Econometrics*, 150(2), 207–218.

Linton, O., J. Pan and H. Wang (2010): Estimation for a Nonstationary Semi-Strong GARCH(1,1) Model with Heavy Tailed Errors, *Econometric Theory*, 26, 1–28.

Ljung, G.M. and G.E.P. Box (1978): On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 65(2), 297–303.

Lucas, R.J. (1976): Econometric Policy Evaluation: A Critique, *Carnegie-Rochester Conference Series on Public Policy*, 1(1), 19–46.

Lundbergh, S. and T. Teräsvirta (2002): Evaluating GARCH Models, *Journal of Econometrics*, 110, 417–435.

Lütkepohl, H. and A. Netsunajev (2014): Structural Vector Autoregressions with Smooth Transition in Variances: The Interaction between U.S. Monetary Policy and the Stock Market, Discussion Papers of DIW Berlin No. 1388, DIW Berlin, German Institute for Economic Research.

Manski, C.F. (2004): Measuring Expectations, *Econometrica*, 72(5), 1329–1376.

Martin, A. and J. Ventura (2011): Theoretical Notes on Bubbles and the Current Crisis, *IMF Economic Review*, 59(1), 6–40.

Martin, A. and J. Ventura (2012): Economic Growth with Bubbles, *American Economic Review*, 102(6), 3033–3058.

McNeil, A.J. and R. Frey (2000): Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach, *Journal of Empirical Finance*, 7(3-4), 271–300.

Nelsen, R.B. (2006): *An Introduction to Copulas*, New York: Springer.

Patelis, A.D. (1997): Stock Return Predictability and The Role of Monetary Policy, *The Journal of Finance*, 52(5), 1951–1972.

Pavlidis, E., A. Yusupova, I. Paya, D. Peel, E. Martinez-Garcia, A. Mack and V. Grossman (2014): Episodes of Exuberance in Housing Markets: In Search of the Smoking Gun, Globalization and Monetary Policy Institute Working Paper No. 165, Federal Reserve Bank of Dallas.

Perron, P. (1988): Trends and Random Walks in Macroeconomic Time Series, *Journal of Economic Dynamics and Control*, 12(2), 297–332.

Phillips, P.C.B., S. Shi and J. Yu (2015): Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500, *International Economic Review*, 56(4), 1043–1078.

Phillips, P.C.B., Y. Wu and J. Yu (2011): Explosive Behavior in the 1990's NASDAQ: When Did Exuberance Escalate Asset Values?, *International Economic Review*, 52(1), 201–226.

Posen, A.S. (2006): Why Central Banks Should Not Burst Bubbles, *International Finance*, 9(1), 109–124.

Prieto, E., S. Eickmeier and M. Marcellino (2016): Time Variation in Macro-Financial Linkages, *Journal of Applied Econometrics*, 31(7), 1215–1233.

Primiceri, G.E. (2005): Time Varying Structural Vector Autoregressions and Monetary Policy, *The Review of Economic Studies*, 72(3), 821–852.

Rigobon, R. and B. Sack (2004): The Impact of Monetary Policy on Asset Prices, *Journal of Monetary Economics*, 51(8), 1553–1575.

Romano, J.P. and M. Wolf (2005): Stepwise Multiple Testing as Formalized Data Snooping, *Econometrica*, 73(4), 1237–1282.

Rossi, B. and T. Sekhposyan (2010): Have Economic Models' Forecasting Performance for US Output Growth and Inflation Changed over Time, and When?, *International Journal of Forecasting*, 26(4), 808–835.

Sarma, M., S. Thomas and A. Shah (2003): Selection of Value-at-Risk Models, *Journal of Forecasting*, 22(4), 337–358.

Scherbina, A. and B. Schlusche (2014): Asset Price Bubbles: A Survey, *Quantitative Finance*, 14(4), 589–604.

Schularick, M. and A.M. Taylor (2012): Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008, *American Economic Review*, 102(2), 1029–1061.

Shiller, R.J. (2005): *Irrational Exuberance*, 2^{nd} Edition, Princeton: Princeton University Press.

Stiglitz, J.E. (1990): Symposium on Bubbles, *Journal of Economic Perspectives*, 4(2), 13–18.

Stock, J.H. and M.W. Watson (1999): Forecasting Inflation, *Journal of Monetary Economics*, 44(2), 293–335.

Stock, J.H. and M.W. Watson (2003): Forecasting Output and Inflation: The Role of Asset Prices, *Journal of Economic Literature*, 41(3), 788–829.

Tibshirani, R. (1996): Regression Shrinkage and Selection via the Lasso, *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267–288.

Timmermann, A. (2006): Forecast Combinations, in: C. G. G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*: Elsevier, Vol. 1, Chapter 4, pp. 135–196.

Trichet, J.C. (2005). Asset Price Bubbles and Monetary Policy. Speech at Bridgewater College, Bridgewater, Virginia.

Ulbricht, D., K.A. Kholodilin and T. Thomas (2016): Do Media Data Help to Predict German Industrial Production?, *Journal of Forecasting*.

Wu, J.C. and F.D. Xia (2014): Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound, Working Paper No. 20117, National Bureau of Economic Research.

Zakoian, J.M. (1994): Threshold Heteroskedastic Models, *Journal of Economic Dynamics and Control*, 18(5), 931–955.

Zhu, D. and J.W. Galbraith (2011): Modeling and Forecasting Expected Shortfall with the Generalized Asymmetric Student-t and Asymmetric Exponential Power Distributions, *Journal of Empirical Finance*, 18(4), 765–778.

Zou, H. and T. Hastie (2005): Regularization and Variable Selection via the Elastic Net, *Journal of the Royal Statistical Society, Series B*, 67, 301–320.

Summary

Throughout history, bursting asset price bubbles have frequently challenged not only the stability of the financial system, but have also caused severe economic contractions. Most recently, the global financial crisis (GFC) resulting from the burst of the U.S. housing bubble has provided a forceful reminder about the risks inherent in financial markets and has challenged the understanding of macro- and financial economists about the linkages between the financial system and the real economy. As a result, the crisis has also sparked intense debates about pre-crisis economic and financial policies, in particular with regard to financial market regulation and the role of monetary policy in amplifying or dampening asset price cycles. This dissertation consists of four chapters that empirically address some of these challenges and debates. Thereby, this thesis contributes to the literature on risk modeling of serially dependent asset returns; the real-time detection of asset price bubbles; forecasting of real economic activity using real-time indicators for asset price bubbles; and the role of monetary policy in asset mispricing.

The first chapter, based on a paper with Helmut Herwartz and Moritz Seidel, explores whether model residuals from the class of (threshold) generalized autoregressive conditional heteroskedasticity ((T)GARCH) models are characterized by serial dependence, which could potentially be used to enhance conventional risk forecasts. We find that these residuals are hardly independent and identically distributed but instead show forms of higher order serial dependence. This suggests that TGARCH models commonly employed for predicting market risk of speculative asset returns do not use all available information for their forecast. We propose two strategies to quantify the serial dependence structures between model innovations, a nonparametric estimation approach and a flexible modeling approach based on standardized copula distributions. We show that these strategies more accurately describe the in-sample dependence patterns between consecutive innovations, and outperform conventional TGARCH model predictions for the conditional Value-at-Risk and the conditional Expected Shortfall at the relevant risk levels outlined by the Basel Committee on Banking Supervision.

The second chapter assesses whether emerging asset price bubbles can be detected in real-time. For this, I begin by evaluating the accuracy of existing early warning indicators for stock price bubbles. I apply these indicators to U.S. stock market data and highlight the considerable signal heterogeneity across all indicators, with frequent false positive signals

during normal times and instable signals during the 1990's dot-com bubble. To improve signal accuracy, I then propose two strategies to combine signals from all individual indicators in real-time. First, I put forward a simple counting approach that requires the number of simultaneous bubble signals from all indicators to exceed a specified threshold. Second, I develop a combination indicator based on a multiple testing procedure that controls the overall size of such a joint test. Through simulations, I show that both combination strategies provide more accurate real-time signals for the emergence and collapse of asset price bubbles.

The third chapter, based on joint work with Dirk Ulbricht, assesses whether real-time indicators for bubbles on stock and housing markets contain valuable information for predicting real economic activity. We find that several indicators for asset price bubbles strongly improve upon an autoregressive (AR) forecast model for output growth. Moreover, these bubble-augmented AR models are also highly competitive in providing accurate forecasts against a large set of 216 models based on macroeconomic and financial predictors commonly used to forecast real economic activity. In addition, we note that the best predictive bubble indicators also provide the most plausible bubble signals, providing further evidence that these indicators are capable of detecting true bubble episodes in real-time.

The fourth chapter, based on joint work with Kerstin Bernoth, investigates the role of monetary policy in misaligning stock prices from their fundamental value. Using a structural vector autoregressive model, we decompose the estimated response of stock prices to a monetary policy shock into a change of expected future dividends and a change in the equity risk premium. We find that only about one third of the overall impact response of stock prices can be accounted for by these two sources, which suggests a strong and systematic overreaction of stock markets to monetary policy shocks. This result lends support to the proponents of an activist, leaning against the wind monetary policy: By raising interest rates, the central bank can indeed lower the mispricing component in stock prices. However, this comes at the cost of dampening real economic activity and is hence only recommendable to an inflation-targeting central bank when a perceived excessive asset price boom is accompanied by economic growth and inflation above the bank's targets.

Keywords: Asset price bubbles, financial stability, macro-financial linkages, monetary policy, asset pricing, risk forecasting, volatility modeling, real-time bubble detection, indicator combination, real-time forecasting, leaning against the wind, stock pricing, housing bubbles, structural vector autoregression, unit root testing, GARCH, copula distributions

JEL Classification: C22, C32, C51, C52, C53, E44, E52, G12, G32

Zusammenfassung

Auf Vermögensmärkten ist es im Laufe der Geschichte immer wieder zu Preisblasen gekommen, welche häufig mit schwerwiegenden Folgen nicht nur für die Stabilität des Finanzsystems, sondern auch mit empfindlichen realwirtschaftlichen Konsequenzen einhergingen. Zuletzt hat die aus dem Platzen der U.S.-amerikanischen Immobilienblase resultierende Globale Finanzkrise die Risiken im Finanzsystem deutlich hervortreten lassen, und das Vorkrisenverständnis von Makroökonomen und Finanzwissenschaftlern über die makrofinanziellen Verknüpfungen herausgefordert. Darüber hinaus hat die Krise Debatten über die wirtschafts- und finanzmarktpolitischen Rahmenbedingungen angestoßen, insbesondere mit Blick auf die Finanzmarktregulierung und die Rolle der Geldpolitik in der Verstärkung von Vermögensblasen. Diese Dissertation besteht aus vier Kapiteln, die sich empirisch mit einigen dieser Debatten befassen. Insbesondere leistet diese Dissertation einen Beitrag zur Literatur der Modellierung von seriell abhängigen Finanzmarktrenditen, der Echtzeiterkennung von Vermögenspreisblasen und deren Nutzen für die Konjunkturprognose, sowie der Rolle der Geldpolitik in der Fehlbepreisung von Vermögenswerten.

Das erste Kapitel, basierend auf einem Papier mit Helmut Herwartz und Moritz Seidel, untersucht, ob Modellinnovationen von *(Threshold) Generalized Autoregressive Conditional Heteroskedasticity* ((T)GARCH) Modellen serielle Abhängigkeiten aufweisen, welche potenziell zur Verbesserung herkömmlicher Risikoprognosen genutzt werden könnten. Wir beobachten, dass diese Residuen in der Tat nicht unabhängig und gleichverteilt sind, sondern seriell abhängig in höherer Ordnung sind. Das deutet darauf hin, dass TGARCH Modelle, welche oftmals zur Prognose von Marktrisiken spekulativer Vermögenspositionen verwendet werden, nicht alle verfügbaren Informationen nutzen. Um diesen Prognosegehalt auszunutzen, schlagen wir zwei Strategien zur Quantifizierung der zeitlichen Abhängigkeitsstrukturen zwischen Modellinnovationen vor, einen nichtparametrischen Ansatz, und einen flexiblen Modellierungsansatz basierend auf standardisierten Copulaverteilungen. Wir zeigen, dass diese Strategien die Abhängigkeitsmuster aufeinanderfolgender Innovationen besser beschreiben, und die Prognosegüte konventioneller TGARCH-Modelle für den bedingten Wert im Risiko (Value at Risk, VaR) und den bedingten erwarteten Fehlbetrag (Expected Shortfall, ES) deutlich verbessern.

Das zweite Kapitel analysiert, ob entstehende Vermögenspreisblasen in Echtzeit erkannt werden können. Dafür evaluiere ich zunächst die Signalqualität existierender Frühwarnindi-

katoren für Aktienpreisblasen. Für U.S. Daten zeige ich die erhebliche Heterogenität der Blasensignale aller Indikatoren auf, und beobachte zahlreiche falsche positive Signale in Zeiten stabiler Preise, und instabile Signale während der 1990'er Dotcom-Blase. Zur Verbesserung der Signalgenauigkeit schlage ich daher zwei Strategien zur Echtzeitkombination der Signale aller einzelnen Indikatoren vor: eine simple Abzählregel, die die Signale aller einzelnen Indikatoren aggregiert und die Überschreitung eines Schwellenwertes erfordert, sowie einen Indikator basierend auf einem multiplen Testansatz, welcher die Gesamtgröße eines solchen gemeinsamen Tests kontrolliert. Im Rahmen einer Simulationsstudie zeige ich, dass beide Kombinationsansätze präzisere Echtzeitsignale für die Entstehung und den Zusammenbruch von Vermögenspreisblasen senden.

Das dritte Kapitel, beruhend auf einer gemeinsamen Arbeit mit Dirk Ulbricht, untersucht anschließend, ob diese Echtzeitindikatoren für Aktien- und Hauspreisblasen wertvolle Informationen für die Konjunkturprognose liefern. Unsere Ergebnisse zeigen, dass mehrere dieser Indikatoren die Prognosen eines autoregressiven (AR) Modells für das Wachstum der Industrieproduktion deutlich verbessern. Des Weiteren sind diese um Blasenindikatoren erhöhten AR Modelle wettbewerbsfähig gegen eine umfangreiche Reihe von 216 Modellen, die auf Informationen von häufig verwendeten makroökonomischen Zeitreihen und Finanzmarktvariablen zugreifen. Wir zeigen weiterhin, dass die besten Prognoseindikatoren ebenfalls die plausibelsten Blasensignale senden, was ein weiterer Hinweis darauf ist, dass diese Indikatoren in der Lage sind Vermögenspreisblasen in Echtzeit zu erkennen.

Das vierte Kapitel, basierend auf einem Papier mit Kerstin Bernoth, untersucht die Rolle der Geldpolitik in der Blasenbildung auf Aktienmärkten. Mit Hilfe eines strukturellen vektorautoregressiven Modells zerlegen wir dafür die Veränderung von Aktienpreisen auf einen geldpolitischen Schock in die Veränderung zukÃ$\frac{1}{4}$nftiger Dividenden und die Veränderung der erwarteten Risikoprämie. Wir beobachten, dass lediglich ein Drittel der Gesamtveränderung von Aktienpreisen auf diese beiden Größen zurückzuführen ist, was auf eine starke und systematische Überreaktion von Aktienmärkten auf geldpolitische Schocks hindeutet. Unsere Ergebnisse unterstützen dabei die Forderungen nach einer aktiven, "leaning against the wind" Geldpolitik: In dem die Zentralbank die Zinsrate erhöht, kann sie übertriebene Fehlbewertungen von Aktienpreisen senken. Allerdings verursacht dies ein deutliches Abschwächen der Konjunktur und kann daher nur in Einklang mit einer auf Preisstabilität ausgerichteten Geldpolitik gebracht werden, wenn eine wahrgenommene Aktienpreisblase mit Wirtschaftswachstum und Inflation oberhalb der Zentralbankziele einhergeht.

Schlagworte: Vermögenspreisblasen, Finanzstabilität, Makrofinanzielle Verknüpfungen, Geldpolitik, Preisbildung auf Vermögensmärkten, Risikoprognosen, Volatilitätsmodellierung, Echzeitblasenerkennung, Indikatorkombination, Echtzeitprognose, "leaning against the wind", Aktienpreise, strukturelle Vektorautoregression, Einheitswurzeltest, GARCH, Copulaverteilung

JEL Klassifikation: C22, C32, C51, C52, C53, E44, E52, G12, G32

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass ich die vorgelegte Dissertation auf Grundlage der angegebenen Quellen und Hilfsmittel selbstständig verfasst habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, sind als solche kenntlich gemacht. Die vorgelegte Dissertation hat weder in der gleichen noch einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule zum Promotionsverfahren vorgelegen.

Benjamin Beckers

Berlin, den 18. Mai 2017

Liste verwendeter Hilfsmittel

- Matlab 8.1.0.604 (R2013a)
  - Optimization Toolbox
  - Financial Toolbox
  - Econometrics Toolbox
  - Statistics Toolbox
- RStudio 1.0.136 basierend auf R 3.3.0
- Eviews 8.0
- Stata 14
- Microsoft Excel
- LaTeX
- Siehe auch Literatur- und Quellenangaben