Essays on Business Cycle Forecasting and Banking Regulation

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

vorgelegt von
Dipl.-Math. B.Sc. Thomas Theobald
aus Wittlich
Düsseldorf, im August 2014

"Gedruckt mit der Genehmigung des Fachbereichs Wirtschaftswissenschaften der Freien Universität Berlin"

Dekan: Univ.-Prof. Dr. Klaus Ruhnke

Erstgutachter: Univ.-Prof. Dr. Dieter Nautz

Zweitgutachter: Prof. Dr. Sven Schreiber

Tag der Disputation: 01.07.2014

Danksagung

Ich danke Sven Schreiber und Dieter Nautz für gute Betreuung, ausreichende Freiheit und kritische Würdigung, ohne die die Erstellung dieser Arbeit nicht möglich gewesen wäre. Weiterhin danke ich der Hans-Böckler-Stiftung und dem Institut für Makroökonomie und Konjunkturforschung (IMK) für die Finanzierung meiner Promotionstätigkeit. Besonderer Dank gilt dem ganzen IMK Team, insbesondere auch ehemaligen Kollegen, für die stete Hilfsbereitschaft, Diskussionsfreude und tolle Arbeitsatmosphäre während meines Doktorats.

Danken möchte ich auch meinem Koautor Christian R. Proaño. Ferner danke ich anonymen Referees, dem Forschungsnetzwerk Makroökonomie und Makropolitik (FMM) und dem Institute for New Economic Thinking (INET), die mir erlaubt haben, mein ökonomisches Denken im Laufe der Jahre zu schärfen. Schließlich danke ich Rory Tews für sprachliche Korrekturen. Ohne die Unterstützung meiner Familie und meiner Freunde wäre diese Arbeit nicht entstanden.

Eigenanteil der Leistung

Diese Dissertation besteht aus drei Arbeitspapieren, von denen eines in Zusammenarbeit mit einem Koautor entstanden ist.

- Proaño, C. R. und Theobald, T. (2014) 'Predicting Recessions with a Composite Real-Time Dynamic Probit Model', akzeptiert zur Veröffentlichung im *International Journal of Forecasting* . Eine frühere Fassung wurde veröffentlicht als NSSR Working Paper 05/12, The New School for Social Research, New York.

- Theobald, T. (2014) 'Markov Switching with Endogenous Number of Regimes and Leading Indicators in a Real-Time Business Cycle Application', akzeptiert zur Veröffentlichung in *The Empirical Economics Letters* . Eine frühere Fassung wurde veröffentlicht als IMK Working Paper 98, Institut für Makroökonomie und Konjunkturforschung, Düsseldorf.

- Theobald, T. (2014) 'Agent-based Risk Management - A Regulatory Approach to Financial Markets', wieder eingereichte Fassung für *Journal of Economic Studies* . Eine frühere Fassung wurde veröffentlicht als IMK Working Paper 95, Institut für Makroökonomie und Konjunkturforschung, Düsseldorf.

Contents

List of Abbreviations

ABM Agent-based Model

BB Bry-Boschan

BDS Brock-Dechert-Scheinkman

BSM Black-Scholes-Merton

CLI Composite Leading Indicators

DESTATIS Deutsches Statistisches Bundesamt

DSGE Dynamic Stochastic General Equilibrium

ECRI Economic Cycle Research Institute

ESRB European Systemic Risk Board

EURIBOR Euro Inter-Bank Offered Rate

GARCH Generalized Autoregressive Conditional Heteroscedasticity

GDP Gross Domestic Product

IMM Internal Model Method

MAE Mean Absolute Error

MS Markov Switching

MSARX Markov Switching Autoregressive Model with Exogenous Variables

MSVAR Markov Switching Vector Autoregressive Model

NBER National Bureau of Economic Research

OECD Organisation for Economic Co-Operation and Development

RMSE Root Mean Squared Error

VaR Value at Risk

List of Tables

List of Figures

General Introduction and Results

Work on this thesis started around the time of the recent financial crisis (2007 - 2009). What has become apparent since then is the practical use of two research questions dealt with in the following: First, as financial crises are usually accompanied by deep recessions that affect the whole economy, how can a business cyclical turning point be detected as early and reliably as possible under these circumstances? Secondly, taking under-regulation as one of the root causes of the last crisis, can agent-based models be efficiently used for the purpose of banking regulation?

In normal growth periods, regardless of whether growth is negative or positive, conventional linear models are able to predict the current and upcoming quarterly GDP with reasonable forecast accuracy, *i.e.* they are more or less able to bridge the publication lag. However, before the crisis, not only did the forecast performance of most leading indicators in such models decline (Drechsel and Scheufele, 2012), but also the degree according to which the data generating process is well described by a linear model. For example, in Germany the Gross Domestic Product (GDP) dropped as far as 4% from one quarter to the next. Furthermore, in the course of deep recessions, linear models are often subject to a premature mean reversion which turns out to be disadvantageous.

Here, it is helpful to turn to a probabilistic approach and to forecast the business cycle turning points by means of the level of the predicted recession probabilities (Schreiber, Theobald, Proaño, Stephan, Rietzler and Detzer, 2012). In particular, non-linear models with two or more states, such as probit models and Markov regime switching models, can achieve relatively high forecast accuracy in such situations. The first two chapters of this thesis present modifications to these methods and fruitfully applies them to the task of timely and accurate predicting of turning points in the business cycle.

In the meantime, the literature has revealed several causes for the financial crisis starting in 2007. These include the low interest rate policy - especially of the Federal Reserve (FED) after the dot-com bubble (FCI-Commission, 2011) -, rising income inequality (Stiglitz, 2012; Rajan, 2010) and growing current account imbalances (Behringer and van Treeck, 2013; Belabed, Theobald and van Treeck, 2013; Kumhof, Ranciere, Lebarz, Richter and Throckmorton, 2012). Without doubt, an insufficient banking regulation (Horn, Joebges, Kamp, Krieger, Sick and Tober, 2009) should also be mentioned, which allowed banks to underestimate the riskiness of financial market instruments. The systemic risk arising from innovative instruments like securitization was particularly significant, but also feedback effects from conventional derivatives markets.

A potential tool for improving the regulation of financial markets could be (so-called) Agent-Based Models (ABM) - in particular agent-based computational models. Such models are primarily directed at representative agent setups as typically employed in Dynamic Stochastic General Equilibrium (DSGE) models. The advantage is that they produce dynamics of diverse and simultaneously interacting agents reflecting behavioral patterns which, on the one hand, are observable in real-world financial markets like herding and contrarian behavior (Park and Sabourian, 2011) and which, on the other hand, lead to price paths that show typical financial data properties like fat tails and volatility cluster (Pagan, 1996). The third chapter of this thesis therefore deals with the compatibility of agent-based computations and (market) risk management as it is currently applied by banks, in line with the Basel regulatory requirements.

The thesis comprises three individual papers which are included as chapters. An appendix contains supplementary material. While the first two chapters focus on the detection of a recession in the aftermath of a financial crisis, the third one covers under-regulation as one of the root causes of the last financial crisis. The main contributions can be summarized as follows:

- **Chapter 1, Predicting Recessions with a Composite Real-Time Dynamic Probit Model:** In this chapter we propose a composite indicator for real-time recession forecasting based on alternative dynamic probit models. For this purpose, we use a large set of monthly macroeconomic and financial leading indicators from the German and U.S. economies. Alternative dynamic probit regressions are specified through automatized general-to-specific as well as specific-to-general lag selection procedures on the basis of slightly different initial sets, and the resulting recession probability forecasts are then combined in order to decrease the volatility of the forecast errors and increase their forecasting accuracy. This procedure not only features good in-sample forecast statistics, but also good out-of-sample performance, as is illustrated using a real-time evaluation exercise.

- **Chapter 2, Markov Switching with Endogenous Number of Regimes and Leading Indicators in a Real-Time Business Cycle Application:** This chapter uses a broad range of macroeconomic and financial indicators in combination with a Markov Switching (MS) framework. The purpose is to predict business cycle turning points based on monthly German real-time data covering the recession and the recovery after the financial crisis. We show how to take advantage of combining single MSARX forecasts with the adjusting of the number of regimes on the real-time path, which both lead to higher forecast accuracy through the non-linearity of the underlying data-generating process. Adjusting the number of regimes implies distinguishing between recessions which are either normal or extraordinary. In fact it turns out that the Markov Switching model can signal quite early whether a conventional recession will occur or whether an economic downturn will be more pronounced.

- **Chapter 3, Agent-based Risk Management - A Regulatory Approach to Financial Markets:** This chapter provides market risk calculation for an equity-based trading portfolio. Instead of relying on the purely stochastic internal model method, which banks currently apply in line with the Basel regulatory requirements, we propose to include also alternative price mechanisms from the financial literature into the regulatory framework. For this purpose a financial market model based on heterogeneous agents is developed, capturing the realistic feature that parts of the investors do not follow the assumption of no arbitrage, but are motivated by behavioral heuristics instead. Although both the standard stochastic as well as the behavioral model are restricted to a calibration including the last 250 trading days, the latter is able to capitalize possible turbulences on financial markets and likewise the well-known phenomenon of excess volatility - even if the last 250 days reflect a non-turbulent market. Thus, including agent-based models in the regulatory framework could create better capital requirements with respect to their level and countercyclicality. This in turn could reduce the extent to which bubbles arise since market participants would have to anticipate comprehensively the costs of such bubbles bursting. Furthermore a market-wide hedge ratio is deduced from the agent-based construction to lower the influence of speculative derivatives.

The overall goal of the thesis at hand is to contribute to a thorough understanding of recession detection, especially in the aftermath of financial crisis, and of under-regulation as one of the root causes of the last financial crisis. I hope that the empirical results and econometric tools (IMK_Konjunkturindikator, 2012, in particular), which have been developed within the scope of this thesis, are of practical use and that they can be productively applied for policy recommendations. I would also hope that this thesis can motivate future research, with suggestions offered at the end of each chapter.

Allgemeine Einleitung und Resultate

Diese Arbeit ist in zeitlicher Nähe zur jüngsten Finanzmarktkrise (2007-2009) entstanden. Sie behandelt daher zwei Forschungsfelder, deren praktischer Nutzen in Zeiten der Krise offenkundig wurde. Da Finanzmarktkrisen in der Regel von tiefen Rezessionen begleitet werden, die die gesamte Ökonomie betreffen, geht es zunächst um die möglichst frühe, zuverlässige und informative Rezessionserkennung. In einem zweiten Teil der Arbeit wird dann das Thema der Bankenregulierung behandelt, da Unterregulierung zweifelsohne eine der Hauptursachen der letzten Finanzmarktkrise darstellt. Im Detail geht es um die Frage, ob sogenannte Agenten basierte Modelle im Interesse des Gemeinwohls zum Zwecke der Bankenregulierung eingesetzt werden können.

In Phasen 'normalen' Wachstums, unabhängig davon ob negativ oder positiv, scheinen konventionelle lineare Modelle im Stande zu sein mit ausreichender Qualität das Bruttoinlandsprodukt (BIP) des laufenden sowie in abgeschwächter Form auch des nächsten Quartals zu prognostizieren. Im Wesentlichen sind diese Modelle also fähig, die sogenannte Veröffentlichungslücke zu schließen. Vor der Krise hat jedoch nicht nur die Vorhersageleistung der meisten Frühindikatoren in solchen Modellen abgenommen (Drechsel and Scheufele, 2012), sondern auch der Grad, mit dem der Daten generierende Prozess wohl beschrieben ist durch ein lineares Modell. Zum Beispiel ist das BIP in Deutschland in der Spitze von einem Quartal zum anderen um mehr als 4% gesunken. Desweiteren leiden lineare Modelle im Verlauf einer tiefen Rezession oft unter einer zu frühzeitigen Mittelwert-Rückkehr, so dass die Tiefe der Rezession nicht korrekt prognostiziert wird.

In diesem Fall wird es hilfreich, sich einem probabilistischen Ansatz zu zuwenden und sich auf die Vorhersage der konjunkturellen Wendepunkte mittels der Höhe der prognostizierten Rezessionswahrscheinlichkeit zu konzentrieren (Schreiber et al., 2012). Insbesondere nicht-lineare Modelle mit zwei oder mehr (konjunkturellen) Zuständen, wie etwa probit Modelle oder Markov Regime

Switching Modelle, können eine relativ hohe Vorhersagegüte in solchen Situationen erreichen. Die ersten beiden Kapitel dieser Arbeit stellen Modifikationen dieser Methoden dar, um diese Gewinn bringend im Hinblick auf ein möglichst frühzeitiges und exaktes Erkennen konjunktureller Wendepunkte einzusetzen.

Zwischenzeitlich hat die Literatur mehrere Ursachen aufgezeigt für das Ausbrechen der Finanzmarktkrise ab 2007. Diese beinhalten unter anderem die anhaltende Niedrigzinspolitik, insbesondere durch die amerikanische Zentralbank (FED) im Nachklang der Dotcom-Blase (FCI-Commission, 2011), zunehmende Einkommensungleichheit (Stiglitz, 2012; Rajan, 2010) und auch daraus resultierende wachsende Leistungsbilanzungleichgewichte (Behringer and van Treeck, 2013; Belabed et al., 2013; Kumhof et al., 2012). Ohne Zweifel muss jedoch auch eine zu laxe Bankenregulierung (Horn et al., 2009) erwähnt werden. Diese erlaubte Banken, den Risikogehalt ihrer Finanzinstrumente zu unterschätzen, insbesondere das systemische Risiko innovativer Instrumente wie Verbriefungen, aber auch Rückkopplungseffekte bis dato schon etablierter Derivatemärkte.

Ein mögliches Instrument, um die Regulierung der Finanzmärkte zu verbessern, können sogenannte Agenten-basierte Modelle sein, insbesondere 'Agent-based Computational Economic' (ACE) Modelle. Solche Modelle wenden sich unmittelbar gegen das Konzept des Repräsentativen Agenten, wie es typischerweise in Dynamischen Stochastischen Allgemeinen Gleichgewichtsmodellen (DSGE) verwendet wird. Ihr Vorteil liegt darin, eine Preisdynamik auf Grundlage verschiedener und gleichzeitig interagierender Agenten zu produzieren. Diese Dynamik spiegeln Verhaltensmuster wider, wie sie auch auf realen Finanzmärkten zu beobachten sind, wie etwa Herden- bzw. gegenläufiges Verhalten (Park and Sabourian, 2011). Zudem weisen hieraus entstehende Preispfade statistische Eigenschaften auf, wie man sie typischerweise von der Untersuchung von Finanzmarktdaten kennt. Hierzu gehören etwa 'fat tails' und Volatilitätscluster (Pagan, 1996). Das dritte Kapitel dieser Arbeit beschäftigt sich daher mit der Kompatibilität von ACE Modellen und (Markt-)risiko Management, wie es Banken derzeit gemäß der in den Basel-Regularien festgeschriebenen Internen Modelle Methode anwenden.

Die vorliegende Arbeit umfasst drei individuelle Papiere, die die Kapitel der Arbeit darstellen. Ein Anhang enthält ergänzendes Material. Während sich die beiden ersten Papiere mit der Rezessionserkennung zu Zeiten der Finanzmarktkrise beschäftigen, behandelt das dritte Papier Unterregulierung als eine der wesentlichen Ursachen der letzten Finanzmarktkrise. Die Haupterkenntnisse können wie folgt zusammengefasst werden:

- **Kapitel 1, Vorhersage von Rezessionen in Echtzeit mit einem zusammengesetzten dynamischen Probitmodell:** In diesem Kapitel schlagen wir einen zusammengesetzten Indikator zur Echtzeit-Rezessionsvorhersage vor - basierend auf unterschiedlichen dynamischen Probitmodellen. Innerhalb des Modell wird eine große Menge sowohl makroökonomischer als auch Finanzmarkt-Frühindikatoren für Deutschland und für die USA verarbeitet. Alternative dynamische Probitregressionen werden zunächst mittels automatisierter Lagauswahlprozeduren auf Basis einer sich unterscheidenden Ursprungsmenge an Variabeln spezifiziert. Die resultierenden prognostizierten Rezessionswahrscheinlichkeiten werden dann aggregiert mit dem Ziel, die Volatiltität der Vorhersagefehler zu senken und die Vorhersagegenauigkeit zu erhöhen. Diese Vorgehensweise führt nicht nur zu guten In-Sample Resultaten, sondern auch zu einer guten Vorhersageleistung Out-of-Sample, wie eine Evaluation unter Echtzeitbedingungen zeigt.

- **Kapitel 2, Markov Switching mit endogener Anzahl von Regimen in Anwendung auf Echtzeit-Konjunkturprognose:** Dieses Kapitel nutzt eine breite Palette von Frühindikatoren im Zusammenspiel mit einem Markov Switching Modell. Zweck ist es, konjunkturelle Wendepunkte voherzusagen - in diesem Kapitel basierend auf monatlichen deutschen Echtzeitdaten, welche die Rezession in Folge der Finanzmarktkrise sowie die anschließenden konjunkturelle Erholung beinhalten. Wir zeigen, wie man Vorteile aus der Kombinationsbildung einzelner MSARX Vorhersagen sowie aus der Echtzeit-Anpassung der Anzahl der Regime ziehen kann. Beides führt durch die Nichtlinearität des Daten generierenden Prozess zu höherer Vorhersagegenauigkeit. Eine Anpassung der Regimeanzahl impliziert eine Unterscheidung zwischen Rezessionen, die (in ihrer Tiefe) als normal oder als außergewöhnlich eingestuft werden. Es zeigt sich, dass das Markov Switching Modell relativ früh signalisieren kann, ob es sich eher um eine konventionelle Rezession handelt oder ob der Rückgang der ökonomischen Aktivität deutlich stärker ausfallen wird.

- **Kapitel 3, Agenten-basiertes Risiko Management - Ein Regulierungsansatz für Finanzmärkte:** Dieses Kapitel behandelt Methoden zur Berechnung des Marktrisikos von (Aktien-basierten) Handelsportfolien. Statt sich ausschließlich auf die stochastische Interne Modelle Methode zu verlassen, wie sie derzeit von Banken gemäß dem Basel Regelwerk angewendet wird, empfiehlt dieses Kapitel auch alternative

Preismechanismen zu berücksichtigen, wie sie in der wissenschaftlichen Finanzliteratur bekannt sind. Zu diesem Zweck wird eine Modifikation des Modells von Lux and Marchesi (2000) entwickelt, um den gegenwärtigen Regulierungsrahmen zu erweitern. Dieses Modell besitzt als Merkmal, dass nicht alle Investoren der No-Arbitrage Annahme folgen, sondern sich vielmehr von Faustregeln und Marktpsychologie leiten lassen. Obwohl beide Ansätze, der standard-stochastische und der verhaltensorientierte, jeweils eine Datenhistorie von 250 Handelstagen nutzen, wie sie der Regulierungsrahmen vorgibt, kann nur der verhaltensorientierte Ansatz, Turbulenzen am Finanzmarkt abbilden - auch wenn die letzten 250 Handelstage einem relativ ruhigen Markt entsprechen, für den die Effizienzmarkthypothese nicht abgelehnt werden kann. Es wird argumentiert, dass es sinnvoll wäre, im regulatorischen Kapital das Maximum der Value-at-Risk (VaR) aus beiden Ansätzen zu verwenden, um umfassendere und stärker anti-zyklische Eigenkapitalanforderungen stellen zu können. Das wiederum könnte dazu führen, dass (irrationale) Preisblasen in deutlich geringerem Ausmaß entstehen, da die Marktteilnehmer das Platzen solcher Blasen über die Eigenkapitalanforderungen kostenmäßig berücksichtigen müssen. Schließlich ergibt sich aus dem vorgestellten Agenten-basierten Modell eine Schlüsselgröße, die helfen kann, die nachteiligen Folgen spekulativer Derivategeschäfte zu vermindern.

Das Gesamtziel der vorliegenden Doktorarbeit ist es, zu einem gründlichen Verständnis von Rezessionsvorhersagen, insbesondere in Folgen von Finanzmarktkrisen, sowie von Unterregulierung als eine der zentralen Ursachen der letzten Finanzmarktkrise beizutragen. Ich hoffe, dass die empirischen Resultate und ökonometrischen Werkzeuge, die im Rahmen dieser Doktorarbeit entwickelt wurden, insbesondere IMK_Konjunkturindikator (2012), von praktischem Nutzen sind und dass sie fruchtbar für Politikempfehlungen eingesetzt werden können. Falls diese Arbeit zukünftige Forschung motivieren kann - Vorschläge hierfür werden jeweils am Ende der Kapitel unterbreitet - würde ich mich sehr freuen.

Chapter 1

Predicting Recessions with a Composite Real-Time Dynamic Probit Model

1.1 Introduction

The timely and accurate prediction of turning points in the business cycle is one of the most policy-relevant aspects of macroeconomic forecasting. This task is, however, also one of the most challenging: Not only are there many potential nonlinearities at the onset of a turning point in economic activity, but also significant uncertainty around macroeconomic data at the current edge[1], as well as the model uncertainty inherent in all applied work.

To mitigate the model uncertainty problem, Bates and Granger (1969) were among the first to propose a combinatorial approach. They showed that the inclusion of inferior ex-ante forecasts could increase the predictive power of the best ex-ante forecasts if they contained some novel information. More recently, Timmermann (2006) also emphasized the usefulness of forecast combinations due to (1) diversification, (2) structural breaks, (3) misspecification of individual forecasts and (4) systematic differences in the individual loss functions.

In contrast, methods for reducing the uncertainty inherent in end-point data are less developed. Pesaran and Timmermann (2005) have stressed the urgent need to develop robust interactive systems of model specification and evaluation designed explicitly to work in real time, as "by setting out in advance a set of

[1] The current edge is defined as the last observation(s) of a certain vintage of macroeconomic data. These observations are usually subject to future data revisions. They are also called end-point data.

rules for observation windows and variable selection, estimation, and modification of the econometric model, automation provides a way to reduce the effects of data snooping and facilitates learning from the performance of a given model when applied to a historical data set" (Pesaran and Timmermann, 2005, p.212).

Binary response models have been used extensively in the literature for the prediction of business cycle turning points (Estrella and Mishkin, 1998; Bernand and Gerlach, 1998; Estrella, Rodrigues and Schich, 2003; Moneta, 2005; Wright, 2006; Haltmaier, 2008; Rudebusch and Williams, 2009; Chen, Iqbal and Lai, 2011; Hao and Ng, 2011; Ng, 2012). Along these lines we discuss the rationale and structure of a composite indicator for real-time recession forecasting based on alternative dynamic probit models specified through automatized *general-to-specific* and *specific-to-general* variable and lag selection procedures. This approach is specifically designed to work under real time conditions as discussed in Proaño (2010).

The main contribution of this chapter is thus the development of a composite dynamic probit indicator along the lines of recent studies using binary response models such as Kauppi and Saikkonen (2008) and Nyberg (2010) for monthly recession forecasting under real-time conditions. As will be discussed in this chapter, the estimation of several dynamic probit regressions and the combination of the resulting recession probability estimates takes into consideration the information of additional leading indicators and achieves a higher recession forecast accuracy.

The remainder of this chapter is organized as follows: In Section 1.2 we discuss in detail the structure of the composite real-time dynamic probit model and its underlying combination scheme. In Section 1.3 the real-time in- and out-of-sample performance of the composite model for the German and U.S. economies are presented. A comprehensive comparison between our model and other existing approaches is conducted in Section 1.4. Finally, Section 1.5 draws some conclusions from this study and points out possible extensions for future research.

1.2 Methodology

Following the work of Estrella and Hardouvelis (1991), binary response models have been widely used for the estimation and forecasting of recessionary periods over the last twenty years (Dueker, 1997; Kauppi and Saikkonen, 2008; Rudebusch and Williams, 2009; Nyberg, 2010). In this strain of the literature, the binary recession indicator series b_t, which represents the state of the economy within the business cycle, is set such that

$$b_t = \begin{cases} 1, \text{ if the economy goes through a recessionary phase at time } t \\ 0, \text{ if the economy experiences an expansion at time } t. \end{cases}$$

Let Ω_{t-h} be the information set available at $t-h$, where h represents the forecasting horizon. Assuming a one-period ahead forecast horizon $h=1$, E_{t-1} and $Pr_{t-1}(\,\cdot\,)$ denote the conditional expectation and the conditional probability given the information set Ω_{t-1}. Under the assumption that b_t has a Bernoulli distribution conditional on Ω_{t-1}, i.e.

$$b_t|\,\Omega_{t-1} \sim \mathcal{B}(p_t),$$

the conditional probability p_t of b_t taking the value 1 in t is given by

$$E_{t-1}(b_t) = Pr_{t-1}(\,b_t = 1) = p_t = \Phi(E(\varphi_{t|t-1})),$$

where φ_t represents a linear combination of the random variables contained in the information set Ω_{t-1}. $\Phi(\,\cdot\,)$ represents the linking function between φ_t and the conditional probability $\Pr_{t-1}(\,b_t=1)$ according to the Bernoulli distribution, which in probit models is given by the standard normal distribution function.

The latent variable of the real-time dynamic probit indicator at hand is explained by various lags of the autoregressive reference series and a set of exogenous macroeconomic and financial leading indicators (which we discuss in detail below) summarized in the matrix $\mathbf{x}_t$, i.e.

$$\varphi_t = \sum_{j=h+D_y}^{p} \alpha_j y_{t-j} + \sum_{j=h+D_{\mathbf{x}}}^{q} \mathbf{x}'_{t-j}\beta_j + u_t, \quad u_t \sim N\,(0,1) \quad \forall t, \tag{1.1}$$

where D_y and D_x stand for the real-time data availability constraints.[2]

It should be clear that the inclusion of a large set of variables in $\mathbf{x}_t$ may lead to a serious multicollinearity problem if some series are highly correlated with others. This is likely to be the case if interest rates of government bonds at different maturities (or their spreads vis-à-vis the short-term interest rate) are

[2]In Proaño (2010) the latent variable is explained by various lags of the lagged binary variable b_t in addition to the lagged reference series, i.e.

$$\begin{aligned} \varphi_t = & \sum_{j=h+R}^{o} \delta_j b_{t-j} + \sum_{j=h+D_y}^{p} \alpha_j y_{t-j} + \sum_{j=h+D_{\mathbf{x}}}^{q} \mathbf{x}'_{t-j}\beta_j + u_t, \\ & u_t \sim N\,(0,1) \quad \forall\, t, \quad R > D_y, \end{aligned} \tag{1.2}$$

where R stands for the recession recognition lag. Although we find that the inclusion of lags with this latter variable slightly improves the out-of-sample real-time forecast accuracy (see Appendix A), the inclusion of both autoregressive terms may produce multicollinearity problems. Following the advice of one referee, we only include the lagged values of the reference series y_t.

included at the same time in $\mathbf{x}_t$. In order to avoid this problem, we consider different specifications represented by $\mathbf{z}_t^i$ (the matrix which contains all the explanatory variables of that particular specification), $i \in I$. Accordingly, an i-th specification of the h-step ahead recession forecast of the probit model regression is given by

$$\begin{aligned} \varphi_{t+h}^i = \mathbf{z_t^i}'\beta + u_{t+h}^i, \quad u_{t+h}^i \sim N(0,1), \quad i \in I, \quad \text{with} \\ b_{t+h}^i = \begin{cases} 1 & : \quad \varphi_{t+h}^i > 0 \\ 0 & : \quad \varphi_{t+h}^i \leq 0 \end{cases} \end{aligned} \tag{1.3}$$

where the size of I is equal to the product of the combinatorial dimension and the elements in each of its components. For instance, with five different interest rate spreads and two different lag selection procedures, ten specifications can be taken into account.

Further, in order to avoid the latent problem of choosing an arbitrary model specification based on an ad-hoc selection of lagged values – and of the explaining variables in general –, each alternative dynamic probit specification is estimated using a *general-to-specific* (G) as well as a *specific-to-general* (S) approach following Proaño (2010). In the *general-to-specific* selection procedure (Campos, Ericsson and Hendry, 2005), the explanatory contribution of each lag of each explanatory variable is tested using a redundant variables Likelihood Ratio (LR) test, with the LR statistic computed as

$$LR = -2(\mathcal{L}_R - \mathcal{L}_U)$$

where $\mathcal{L}_\mathcal{R}$ and $\mathcal{L}_\mathcal{U}$ are the maximized values of the (Gaussian) log likelihood function of the unrestricted and restricted regressions.[3] In the *specific-to-general* selection procedure, in contrast, the added explanatory value of an additional lag of each explanatory variable was tested using an omitted variables Likelihood Ratio test, where under the H_o the coefficient of the additionally-added variable (lag) is not significant.

Given the uncertainty linked with the use of macroeconomic data, as well as the potential misspecification of some/all of the alternative dynamic probit regressions, it is impossible both a-priori and a-posteriori to select one particular specification as "the one" best data-generating process. Accordingly, we pursue a combinatorial approach where the information of each regression is incorporated while its eventual bias is balanced.

To express such an approach, following Theobald (2012) let

$$\mu_{t+h|t} = \left(\mu_{t+h|t}^1, \ldots, \mu_{t+h|t}^{|I|}\right)'$$

[3] Under the H_o of this asymptotically χ^2 distributed test with one degree of freedom, the coefficient of a redundant variable (lag) is zero. A rejection of this test results in the tested variable (lag) remaining in the model specification.

denote the vector of single forecasts and

$$\theta = C\left(\mu_{t+h|t}; w_c\right)$$

represent the combinatorial forecast resulting from the aggregation of the underlying forecasts by means of determinate combination weights. In the simple case of equally-weighted recession probability forecasts, the combinatorial forecast would then be given by[4]

$$\theta = \frac{1}{|I|}\sum_{i=1}^{|I|}\mu^i_{t+h|t}, \qquad |I| = \#\{\text{different regressor set}\} \times \#\{G, S\}. \tag{1.4}$$

In the present paper, the difference of initial regressor sets is generated by considering five long-term maturities (1, 2, 3, 5 and 10 years), where for each the corresponding spread is calculated by subtracting the 3 month euribor interest rate. Hence, one obtains $|I| = 10$ for the *simple* average approach. In terms of the expected future value conditional on current information the combinatorial approach leads to

$$\begin{aligned} E\left(b^i_{t+h}|\mathbf{z}^i_t, \beta\right) &= \mu^i_{t+h|t} = Pr\left(b^i_{t+h} = 1|\mathbf{z}^i_t, \beta\right) \\ &= \Phi\left(\mathbf{z_t^i}'\beta\right) = \Phi\left(E\left(\varphi^i_{t+h|t}\right)\right). \end{aligned} \tag{1.5}$$

At this point we could formulate more sophisticated pooling operators for the presented probit models. Based on the analysis in Theobald (2012), however, we formulate a combination scheme which centers the forecast combinations of single forecasts arising from different specifications and model selection procedures around their median and adds the different forecast horizons for the same future value as an additional means of generating the underlying forecasts. This combination scheme can be thought of as a two-stage procedure, where

$$\begin{aligned} \theta &= \sum_{i=1}^{\#\{\text{horizons}\}} \frac{2^{h-1}}{\sum_{j=0}^{\#\{\text{horizons}\}-1} 2^j}\theta^*_h, \; h \in I\backslash I^*, \\ |I| &= \#\{\text{yield spread maturities}\} \times \#\{G, S\} \times \#\{\text{horizons}\}, \end{aligned} \tag{1.6}$$

and

$$\theta^*_k = \sum_{i=1}^{|I^*|} \frac{\left(\sum_{j=1}^{|I^*|} |\mu^{j,k}_{t+h|t} - \mu^{med}_{t+h|t}|\right) - |\mu^{i,k}_{t+h|t} - \mu^{med}_{t+h|t}|}{\left(|I^*| - 1\right)\sum_{j=1}^{|I^*|} |\mu^{j,k}_{t+h|t} - \mu^{med}_{t+h|t}|}, \; i \in I^* \subset I, \; k \in I\backslash I^*. \tag{1.7}$$

med denotes the median of the forecast vector and $I\backslash I^*$ the well-defined set of already aggregated forecasts using the spreads and specification order, while these forecasts still differ in their forecast horizon.

[4]Obviously this is a special case of the *linear opinion pool* with non-negative weights summing up to one.

In real-time applications, including the horizon as an additional generator of the underlying forecasts in eq.(1.6) simply means adjusting the actual prediction by using one forecast which is generated with a longer horizon out of the first revisions and which does not consider the unrevised information of the last observation. In such a constellation it still seems preferable to put more weight on the forecasts which use the most recent available observations, although these are also subject to strong revision. Obviously the number of horizons that can be taken into account is limited by future uncertainty. Aiming at a limited running time of the procedure and taking I^* to have the same size as in the simple average approach, it is reasonable to choose $\#\{\text{horizons}\} = 2$. This leads to $|I| = 20$ underlying forecasts for each prediction generated by what we call the *weighted average* approach.[5] The following section describes the application of this approach to recession forecasting using German and U.S. macroeconomic data.

1.3 Empirical Results

1.3.1 Real-Time Prediction of German Recessions

Data Description

For the following forecasting exercise we employ a wide dataset of German macroeconomic indicators. All financial and real economy variables stem from the Bundesbank database (`www.bundesbank.de/statistik/`), with the exception of the new orders in manufacturing, which stem from the GENESIS-Online database from the German Statistical Office (`www-genesis.destatis.de`), and the ifo business cycle climate index (`http://www.cesifo-group.de`). The estimation sample comprises monthly observations from 1991:1 to 2004:5 (in-sample period) and from 2004:6 to 2011:8 (real-time out-of-sample path).[6]

The binary recession series b_t was computed using the industrial production index as the business cycle reference series as done in a number of previous studies (Anas, M. Billio and Mazzi, 2008; Darne and Ferrara, 2009). Indeed, as discussed in Fritsche and Stephan (2002, p.291), the use of the index of industrial production as a proxy for business cycle movements can be justified

[5]Theobald (2012) also considers a *Bayesian average* approach that is based on the correlations between the forecast errors. However, as found in previous empirical studies (Clemen, 1989; Timmermann, 2006), the results observed in Theobald (2012) indicate that simpler approaches omitting the correlations perform better.

[6]The structural break linked to German reunification limits the length of the data sample considered. As the robustness of maximum likelihood methods depends significantly on the number of observations taken into account in the estimation, we cannot shorten too much our estimation sample to extend the real-time out-of-sample path. This leads to our out-of-sample analysis only being based on one recession period in the German case, in contrast to our analysis of the U.S. economy below.

in that industrial production is "much closer to the 'volatile' aggregates of GDP like investment and exports – which are at the heart of most business cycle theories". Furthermore, the index of industrial production is published on a monthly basis, which not only enhances the timely appraisal of the business cycle, but is also less prone to revisions than the (quarterly) GDP figures.[7]

Specifically, we employ a modified version of the Bry and Boschan (1971) algorithm, according to which a peak in the business cycle is identified when

$$\{y_{t-k} < y_t > y_{t+k}, \quad k = 1, \ldots, 5\}$$

while, analogously, a trough is assumed to take place when

$$\{y_{t-k} > y_t < y_{t+k}, \quad k = 1, \ldots, 5\},$$

where y_t is the two-month moving average of the German index of industrial production – the business cycle reference series.[8]

As an additional censoring rule for the identification of recessionary periods and thus for the generation of the binary recession indicator series b_t, following Harding and Pagan (2002), a *triangle approximation to the cumulative movements* was pursued in order to measure the "severity" of an economic downturn j - and by extension the eventual occurrence of a recession - defined as

$$S_j = 0.5 \times Deepness_j \times Duration_j,$$

where the *duration* is equal to the number of months between peak and trough (according to the NBER's definition, a recession is a *significant decline in economic activity* [of] *more than a few months*), and the *deepness* is defined as the percentage decline,

$$Deepness_j = (y_p - y_t) / y_p,$$

where y_p and y_t are the respective values of the index of industrial production at the corresponding peak and trough, see Anas et al. (2008). A recessionary

[7]As previously discussed, and as pointed out by Burns and Mitchell (1946), an economic recession is characterized by a *widespread* and *synchronized* downturn in overall economic activity observable on a broad set of economic variables. The proper dating of economic expansions and recessions should therefore result from a multivariate approach which takes this into account. For the sake of simplicity and in order to assess the occurrence of turning points on a monthly basis and thus in a more timely fashion, we prefer a univariate business cycle dating approach with the index of industrial production as the business cycle reference series.

[8]Given the high volatility of monthly data, it is usual in the turning points dating literature to "smooth" the underlying business cycle reference series to avoid potential outlier biases. But it is also reasonable not to smooth out too much information. Therefore a low level of smoothing was selected. To check the robustness of our results, we computed the binary recession series using a higher degree of smoothing of the industrial production index. When a three-month moving average was applied to this series, all other dating results are exactly the same, with the following two exceptions: (1) the recessionary period 2003M03-2003M09 changed into 2002M09-2003M09 and (2) the period 2004M07-2004M11 was no longer classified as recessionary. See also the out-of-sample results in the Appendix.

period was identified under the condition $S_j > 0.025$,[9] as there is no universal consensus on the reference minimum duration and deepness of recessions (Darne and Ferrara, 2009, p.5). $S_j > 0.025$ reflects a decline of 1% of the peak level of economic activity coinciding with a duration of at least 5 months.

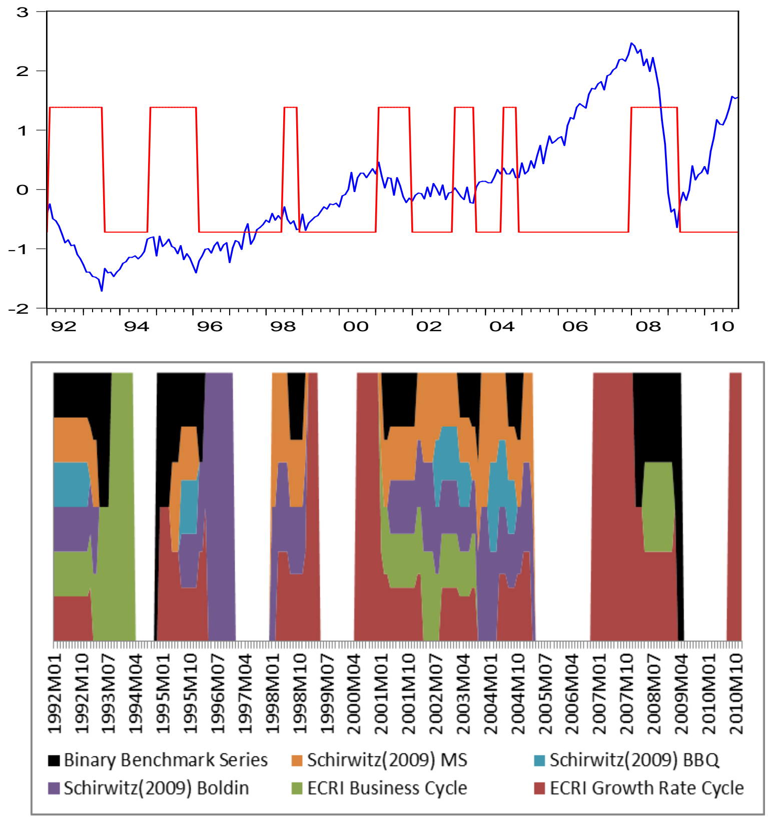


Figure 1.1: Upper graph: German industrial production, business cycle peaks and troughs calculated on the basis of a modified BB algorithm, and related binary recession indicator. Normalized scale. Lower graph: Comparison between our binary recession series (black) and other business cycle chronologies discussed in related studies. The black areas in the lower graph correspond to the red line in the upper graph.

The upper graph in Figure 1.1 illustrates the relationship between the underlying industrial production series and the resulting binary recession indicator series generated by the Bry and Boschan (1971) algorithm. As illustrated there, the implemented algorithm identifies seven recessionary periods for Germany

[9] We also explored the recession dating results under a higher threshold, *i.e.* $S_j > 0.05$. As a consequence, the period 2004M07-2004M11 was no longer classified as recessionary. All other dating results remained the same.

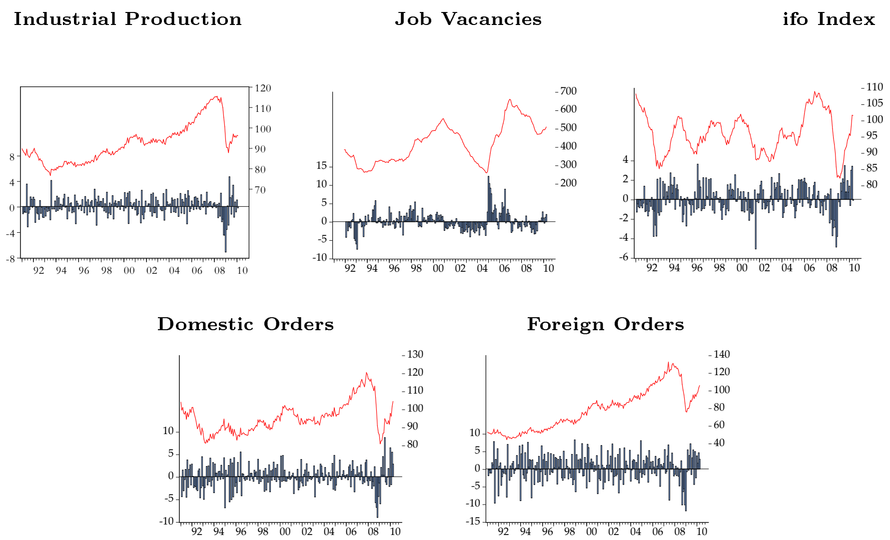


Figure 1.2: Macroeconomic Indicators: Industrial production index, job vacancies, foreign and domestic orders received by the productive sector, and ifo business sentiment index. Sources: Deutsche Bundesbank, DESTATIS, ifo Institute.

between 1991:1 and 2011:8. This number may sound too large, but the following three issues have to be taken into account: First, since the presented probit model cannot differentiate three business cycle regimes (recession, stagnation and expansion), it seems even advantageous that periods of sustainable economic stagnation are detected as recessionary phases, as a missed declaration of a recessionary phase represents the alpha error. Second, due to the limited data availability linked with the major structural break resulting from German reunification, a real-time estimation of a binary series with a lower variation (like the ECRI business cycle) including a comprehensive number of regressors would be much more difficult or even impossible due to the probably occurrence of numerical convergence problems. And finally, in the U.S. case - where it is possible to estimate longer time series and where an official and commonly accepted business cycle dating chronology is available - Figure 1.6 reveals a surprisingly high correlation between our binary recession series and the official NBER announcements. This correlation is not trivial since the ex-post dating algorithm with its fixed data availability (publication lag of the reference series + number of observations necessary for the detection of local extrema) is usually already working ahead of the NBER announcements. For instance, the June 2009 trough was announced in September 2010 (15 months later) whereby the fixed data availability lag is at its maximum equal to 7 months.[10]

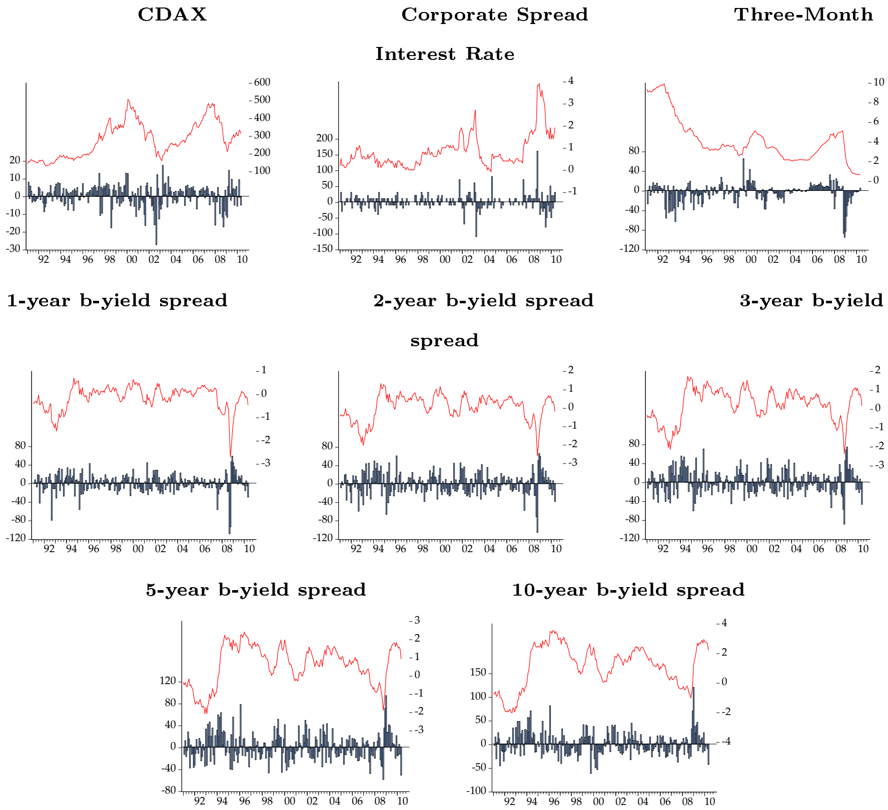


Figure 1.3: Financial Indicators: CDAX, Corporate Spread, Three-Month Euribor and yield spreads of different maturities. Source: Deutsche Bundesbank.

Additionally, we compare our Bry Boschan binary recession series with other dating chronologies from the literature. As can be seen on the lower graph in Figure 1.1, the non-parametric algorithm delivers convincing results (black) as for each of its recession detections at least two other methods from the literature come to the same conclusion. With the business cycle phases from Figure 1.1 in mind, Figures 1.2 and 1.3 show the apparently heterogenous behavior of selected leading indicators during recessionary and expansionary periods. This strongly supports the use of a broad range of regressors. Thus, for the empirical analysis in this chapter, a variety of macroeconomic and financial variables were considered.

[10] Of course, the ex-post dating results can slightly change through data revisions but most of the revisions only take place at the current edge.

Concerning the subset of variables which are supposed to reflect real economic development (besides the index of industrial production), the following indicators were chosen: the open vacancies in the productive sector, the domestic and foreign orders received by the industrial sector, and the ifo business sentiment indicator (all variables given as month-to-month % changes).

The financial indicators were selected to cover all relevant financial market variables. Following Bernanke (1990) and Friedman and Kuttner (1992) the spread between average corporate bond yields of all maturities traded and the average yield of public securities was used, as well as the growth rate of the CDAX price index in order to incorporate German stock market developments. Furthermore, along the lines of Stock and Watson (1989), Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), and more recently Kauppi and Saikkonen (2008) and Nyberg (2010), the yield spread between the long-term and the short-term interest rate – was included in the general set of regressors. More specifically, alternative dynamic probit specifications using the 1-, 2-, 3-, 5- and 10-year yield (calculated by the Svensson's method) spreads as against the three-month EURIBOR were estimated in order to address the uncertainty about which yield spread has the "best" predictive power. The short-term interest rate was also included in the set of regressors. In this respect, Ang, Piazzesi and Wei (2006) show using a dynamic factor model that the two principal factors of the term structure at all traded maturities, which in their study account for 90% of the variation of the whole term structure, are highly correlated with the short-term interest rate and the 10-year yield spread. Additionally, Wright (2006) shows that probit models with the yield spread of the 10-year T-bond to the three-month T-bill *and* the short-term three-month T-bill interest rate outperform probit specifications using only the yield spread in the mean squared error (MSE) sense.

As previously discussed, in order to avoid eventual multicollinearity problems due to the strong correlation between the yield spreads of different maturities, we specified and estimated alternative dynamic probit models underlying the (invariant) set of explaining variables given by the industrial production index, the job vacancies, the ifo business sentiment index, the CDAX price index (all in % month-to-month changes), the corporate spread and the three-month euribor, and alternatively the 1-, 2-, 3-, 5- and 10-year bond yield spreads (to the three-month euribor).[11] In the following, the empirical results of such an automatized model specification procedure are discussed.

[11] In Appendix C we briefly discuss the inclusion of the spread of German securities with respect to U.S. securities. As we discuss there, both the foreign (U.S.) and the domestic term spreads are persistently statistically significant, confirming the previous findings by Bernand and Gerlach (1998) and Nyberg (2010). However, as the inclusion of both series in the same set of regressors creates some multicollinearity problems, we consider only domestic yield spreads.

In-Sample Evaluation

Here we discuss the in-sample results of the dynamic probit specifications obtained by the *general-to-specific* (denoted by a G) as well as the *specific-to-general* (denoted by a S) approaches for 1-, 2- and 3-month ahead forecasts for the estimation sample 1991:1–2010:5. It should be noted, however, that these estimation results represent only an arbitrary "snap-shot" of the performance of the composite indicator, as all regressions underlying it are re-estimated in each and every month based on the newly available information through the automatized real-time specification procedure previously discussed. The estimation results are summarized in tables on page 13 to 15.

A variety of issues are worth highlighting: First, at a general level, the heterogeneity of the dynamic probit model estimations at all three analyzed forecast horizons corroborates the combinatorial approach pursued in this chapter. Indeed, as can be clearly observed in tables on page 13 to 15, the significance level of the majority of variables (lags) is affected by the specific yield spread included in the respective regression sets, as well as by the lag selection procedure (*general-to-specific* or *specific-to-general*). There is, however, a certain "constancy" in the significance level of some variables (lags), which depends on the underlying forecast horizon of the respective regressions. At the one-month-ahead forecast horizon, for example, the sixth lag (relative to the end-point) of the domestic orders and the second to fourth lags of the foreign orders are statistically significant across all probit specifications.

In the same way, the ifo business sentiment index and the lagged reference series do not seem to have any statistical significance at the one-month and the two-month ahead horizon when included among the sets of indicators; the same applies for the job vacancies (at all horizons). But, as we are analyzing a "snap-shot", this does not mean that the ifo business sentiment index, the lagged reference series and the job vacancies do not have any predictive power in general. Actually, the opposite is true for publications (vintages) other than 2010:5, where we find various lags of these variables to be highly significant. Indeed, one advantage of our model is its ability to illustrate that the explanatory power of certain regressors strongly depends on the publication considered. It is also worth mentioning that the statistical significance of the corporate spread series and of the different yield curves seems to be affected by the variables (lag) selection procedures.

As with the one-month-ahead forecast regressions, in the two-month-ahead forecast regressions (at least one lag of) the CDAX price index is statistically significant in almost all specifications, while the importance of the corporate spread and the EURIBOR interest rate seem to increase with the longer horizon.

Summary of Dynamic Probit Regressions, One-Month Forecast Horizon

	Sample: 1991:1 – 2010:5									
	EQ.B-SPRD1YC	EQ.B-SPRD1YS	EQ.B-SPRD2YC	EQ.B-SPRD2YS	EQ.B-SPRD3YC	EQ.B-SPRD3YS	EQ.B-SPRD5YC	EQ.B-SPRD5YS	EQ.B-SPRD10YC	EQ.B-SPRD10YS
IPIDX	-	-	-	-	-	-	-	-	-	-
DOM.ORDERS	6	6	6	6, 7	6, 7	6, 7	6, 7	6, 7	6, 7	6, 7
FOR.ORDERS	2, 3, 4, 5, 6	2, 3, 4, 5	2, 3, 4, 5, 6	2, 3, 4, 5, 6, 7	2, 3, 4, 5, 6, 7	2, 3, 4, 5, 6, 7	2, 3, 4, 5, 6, 7	2, 3, 4, 5, 6, 7	2, 3, 4, 5, 6, 7	2, 3, 4, 5, 6, 7
JOB VAC	-	-	-	-	-	-	-	-	-	-
IFO-IDX	-	-	-	-	-	-	-	-	-	-
CRP_SPRD	-	-	-	-	-	-	-	6	-	6
CDAX	2	-	-	2	2	2	2	2	2	2
EURIBOR 3M	-	-	-	-	1	-	1	6	1	6
B-SPRD1Y	1	1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD2Y	n.a.	n.a.	1	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD3Y	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.
B-SPRD5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	6	n.a.	n.a.
B-SPRD10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	6
SSR	25.135	23.768	25.259	26.359	23.116	26.359	23.116	21.248	23.116	26.085
Avg. Log-Likelihood	-0.387	-0.375	-0.391	-0.396	-0.343	-0.396	-0.343	-0.300	-0.343	-0.309
AIC	0.874	0.868	0.883	0.910	0.787	0.910	0.787	0.718	0.787	0.736
SC	1.044	1.069	1.053	1.111	0.957	1.111	0.957	0.919	0.957	0.937
HQC	0.943	0.950	0.951	0.991	0.856	0.991	0.856	0.799	0.856	0.817
MAE	0.243	0.233	0.246	0.251	0.216	0.251	0.216	0.191	0.216	0.196
RMSE	0.339	0.329	0.340	0.347	0.325	0.347	0.325	0.311	0.325	0.315
Theil	0.339	0.328	0.341	0.350	0.319	0.350	0.319	0.302	0.319	0.306

The upper part of the table shows the results for the automatized general-to-specific and specific-to-general lag selection procedure from the available data at the end of May 2010, while predictions are made for June 2010 (1M forecast horizon). The table displays row-wise the potential regressors and column-wise 10 out of the 20 single specifications which are used for the combinatorial forecast.'-' means that the coefficient of the explanatory variable is not statistically significant at a 5% level. The potential minimum lag of each regressor corresponds to its data availability lag plus the forecast horizon. For instance, '2' means that the coefficient of the variable two months ago is significant for the current binary value at 5%. The lower part shows goodness-of-fit measures for the probit estimations. Specific results such as the overall significance of job vacancies are highlighted in the main text. These results are particularly linked to the available data and thus may change for other publications (vintages).

Summary of Dynamic Probit Regressions, Two-Month Forecast Horizon

	Sample: 1991:1 – 2010:5									
	EQ.B-SPRD1YC	EQ.B-SPRD1YS	EQ.B-SPRD2YC	EQ.B-SPRD2YS	EQ.B-SPRD3YC	EQ.B-SPRD3YS	EQ.B-SPRD5YC	EQ.B-SPRD5YS	EQ.B-SPRD10YC	EQ.B-SPRD10YS
IPIDX	-	-	-	-	-	-	-	-	-	-
DOM.ORDERS	6, 7	6, 7	6, 7	6, 7	6, 7, 8	6, 7, 8	5, 6, 7, 8	6, 7, 8	5, 6, 7, 8	5, 6, 7, 8
FOR.ORDERS	3, 4, 5, 6	3, 4, 5, 6, 7	3, 4, 5, 6	3, 4, 5, 6, 7	3, 4, 5, 6, 7, 8	3, 4, 5, 6, 7	3, 4, 5, 6, 7, 8	3, 4, 5, 6, 7	3, 4, 5, 6, 7, 8	3, 4, 5, 6, 7, 8
JOB VAC	-	-	-	-	-	-	-	-	-	-
IFO-IDX	-	-	-	-	-	-	-	-	-	-
CRP_SPRD	-	-	-	-	6	2	6	-	6	7
CDAX	2	2	2	2	2	2	2	-	2	2
EURIBOR 3M	-	5	-	2	2, 3	2	2, 3	3	3	2, 3
B-SPRD1Y	2	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD2Y	n.a.	n.a.	2	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD3Y	n.a.	n.a.	n.a.	n.a.	2	-	n.a.	n.a.	n.a.	n.a.
B-SPRD5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	2	-	n.a.	n.a.
B-SPRD10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	2	4
SSR	27.423	26.329	27.600	25.691	22.067	25.487	20.732	21.066	21.190	20.839
Avg. Log-Likelihood	-0.415	-0.380	-0.421	-0.372	-0.313	-0.363	-0.297	-0.320	-0.303	-0.299
AIC	0.930	0.852	0.942	0.846	0.764	0.837	0.741	0.787	0.744	0.745
SC	1.101	1.007	1.113	1.017	0.997	1.023	0.989	1.035	0.977	0.993
HQC	0.999	0.914	1.011	0.915	0.858	0.912	0.841	0.887	0.838	0.845
MAE	0.262	0.242	0.266	0.236	0.201	0.232	0.190	0.202	0.195	0.192
RMSE	0.355	0.347	0.356	0.343	0.318	0.342	0.308	0.311	0.312	0.309
Theil	0.361	0.346	0.363	0.343	0.312	0.341	0.300	0.306	0.304	0.301

The upper part of the table shows the results for the automatized general-to-specific and specific-to-general lag selection procedure from the available data at the end of May 2010, while predictions are made for July 2010 (2M forecast horizon). The table displays row-wise the potential regressors and column-wise 10 out of the 20 single specifications which are used for the combinatorial forecast.'-' means that the coefficient of the explanatory variable is not statistically significant at a 5% level. The potential minimum lag of each regressor corresponds to its data availability lag plus the forecast horizon. For instance, '3' means that the coefficient of the variable three months ago is significant for the current binary value at 5%. The lower part shows goodness-of-fit measures for the probit estimations. Specific results such as the overall significance of CDAX are highlighted in the main text. These results are particularly linked to the available data and thus may change for other publications (vintages).

Summary of Dynamic Probit Regressions, Three-Month Forecast Horizon

	Sample: 1991:1 – 2010:5									
	EQ.B-SPRD1YC	EQ.B-SPRD1YS	EQ.B-SPRD2YC	EQ.B-SPRD2YS	EQ.B-SPRD3YC	EQ.B-SPRD3YS	EQ.B-SPRD5YC	EQ.B-SPRD5YS	EQ.B-SPRD10YC	EQ.B-SPRD10YS
IPIDX	-	-	-	-	-	-	-	-	8, 9, 10	-
DOM.ORDERS	6, 7, 8, 9	6, 7, 8	6, 7, 8, 9	6, 7, 8	6, 7, 8,9	6, 7, 8	6, 7, 8,9	6, 7, 8, 9	6, 7	6, 7, 8, 9
FOR.ORDERS	4, 5, 6, 7, 8, 9	4, 5, 6, 7	4, 5, 6, 7, 8	4, 5, 6, 7	4, 5, 6, 7, 8	4, 5, 6, 7	4, 5, 6, 7, 8, 9	4, 5, 6, 7, 8, 9	7, 8, 9	4, 5, 6, 7, 8, 9
JOB VAC	-	-	-	-	-	-	-	-	-	-
IFO-IDX	-	-	-	-	-	-	-	-	-	3
CRP_SPRD	3	-	-	-	-	-	3	-	3	-
CDAX	-	3, 6	3	3	3	3	-	3	8	8
EURIBOR 3M	6	-	-	-	1	-	3	-	4	-
B-SPRD1Y	-	3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD2Y	n.a.	n.a.	3	4	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD3Y	n.a.	n.a.	n.a.	n.a.	3	4	n.a.	n.a.	n.a.	n.a.
B-SPRD5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	3	4	n.a.	n.a.
B-SPRD10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	8	8
SSR	27.186	31.548	31.677	33.045	31.546	32.767	25.554	31.848	23.558	30.969
Avg. Log-Likelihood	-0.371	-0.454	-0.456	-0.475	-0.455	-0.472	-0.355	-0.455	-0.337	-0.432
AIC	0.862	1.019	1.032	1.061	1.029	1.054	0.840	1.030	0.793	0.992
SC	1.064	1.206	1.235	1.248	1.232	1.241	1.058	1.233	0.997	1.210
HQC	0.943	1.095	1.114	1.137	1.111	1.130	0.928	1.112	0.876	1.080
MAE	0.243	0.293	0.295	0.309	0.294	0.307	0.231	0.296	0.217	0.282
RMSE	0.354	0.381	0.382	0.390	0.381	0.389	0.343	0.383	0.330	0.378
Theil	0.354	0.394	0.395	0.408	0.394	0.405	0.341	0.396	0.326	0.386

The upper part of the table shows the results for the automatized general-to-specific and specific-to-general lag selection procedure from the available data at the end of May 2010, while predictions are made for August 2010 (3M forecast horizon). The table displays row-wise the potential regressors and column-wise 10 out of the 20 single specifications which are used for the combinatorial forecast.'-' means that the coefficient of the explanatory variable is not statistically significant at a 5% level. The potential minimum lag of each regressor corresponds to its data availability lag plus the forecast horizon. For instance, '4' means that the coefficient of the variable four months ago is significant for the current binary value at 5%. The lower part shows goodness-of-fit measures for the probit estimations. Specific results such as the overall insignificance of the yield spread are highlighted in the main text. These results are particularly linked to the available data and thus may change for other publications (vintages).

Table 1.1: Expectation-Prediction Evaluations of Probit Regressions, Sample: 1991:1–2010:5

	Expansion	Correct	% Correct	% Incorrect	Recession	Correct	% Correct	% Incorrect
One-Month-Ahead Forecast Horizon								
EQ.B-SPRD1YG	151	138	91.39	8.61	69	36	52.17	47.83
EQ.B-SPRD1YS	151	139	92.05	7.95	69	45	65.22	34.78
EQ.B-SPRD2YG	151	141	93.38	6.62	69	43	62.32	37.68
EQ.B-SPRD2YS	151	139	92.05	7.95	69	38	55.07	44.93
EQ.B-SPRD3YG	151	141	93.38	6.62	69	45	65.22	34.78
EQ.B-SPRD3YS	151	139	92.05	7.95	69	42	60.87	39.13
EQ.B-SPRD5YG	151	138	91.39	8.61	69	47	68.12	31.88
EQ.B-SPRD5YS	151	141	93.38	6.62	69	36	52.47	47.83
EQ.B-SPRD10YG	151	136	90.07	9.93	69	45	65.22	34.78
EQ.B-SPRD10YS	151	135	89.40	10.60	69	46	66.67	33.33
Two-Month-Ahead Forecast Horizon								
	Expansion	Correct	% Correct	% Incorrect	Recession	Correct	% Correct	% Incorrect
EQ.B-SPRD1YG	157	145	92.36	7.64	69	40	57.97	42.03
EQ.B-SPRD1YS	158	143	90.51	9.49	69	36	52.17	47.83
EQ.B-SPRD2YG	157	144	91.72	8.28	69	42	60.87	39.13
EQ.B-SPRD2YS	158	143	90.51	9.49	69	40	57.97	42.03
EQ.B-SPRD3YG	157	144	91.72	8.28	69	41	59.42	40.58
EQ.B-SPRD3YS	160	144	90.00	10.00	69	47	68.12	83.41
EQ.B-SPRD5YG	157	143	91.08	8.92	69	39	56.52	43.48
EQ.B-SPRD5YS	159	141	88.68	11.32	69	45	65.22	34.78
EQ.B-SPRD10YG	158	144	91.14	8.86	69	41	59.42	40.58
EQ.B-SPRD10YS	158	144	91.14	8.86	69	42	60.87	39.13
Three-Month-Ahead Forecast Horizon								
	Expansion	Correct	% Correct	% Incorrect	Recession	Correct	% Correct	% Incorrect
EQ.B-SPRD1YG	158	145	91.77	8.23	69	50	72.46	27.54
EQ.B-SPRD1YS	158	149	94.30	5.70	69	50	72.46	27.54
EQ.B-SPRD2YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD2YS	158	146	92.41	7.59	69	43	62.32	37.68
EQ.B-SPRD3YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD3YS	158	146	92.41	7.59	69	43	62.32	37.68
EQ.B-SPRD5YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD5YS	158	141	89.24	10.76	69	44	63.77	36.23
EQ.B-SPRD10YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD10YS	158	146	92.41	7.59	69	40	57.97	42.03

In-sample Fitted Values

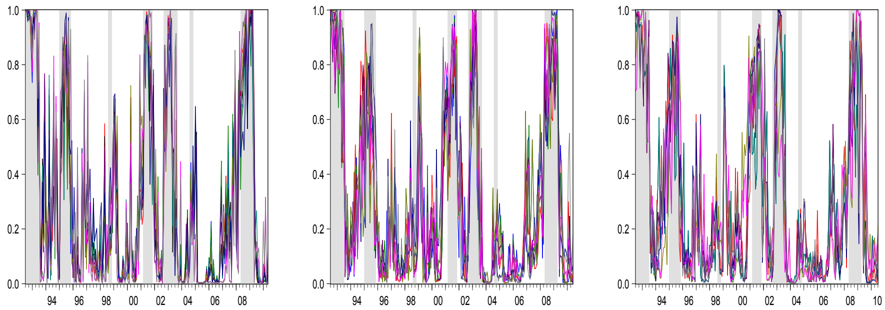


Figure 1.4: In-Sample Fit of Estimated and Average Recession Probabilities, One-Two- and Three-Month Forecast Horizon: As discussed in section 2, different specifications - whose in-sample fits are illustrated here - are used to form a combinatorial forecast, see also Section 1.3.1.

Both domestic and foreign orders (the variables with the highest statistical "constancy" in the previous case) keep their predictive power at the two-month-ahead forecast horizon. Note again, this does not mean that this is always true, as we find other publications where sentiment and financial market variables possess a particularly high explanatory power. Finally, the most remarkable fact concerning the estimation results of the three-month-ahead forecast specifications is that, at least in one specification, the lagged reference series and the ifo business sentiment index turn out to be statistically significant. Moreover, the importance of the yield spread seems to increase slightly as it becomes statistically significant in nine out of ten specifications.

When compared with the outcomes of previous related empirical studies, two of our estimation results are of particular interest: the corroboration of the predictive power of stock price developments vis-à-vis future economic activity already documented by Harvey (1989), Stock and Watson (1999), and recently by Haltmaier (2008); and the finding that the predictive power of the yield spread (irrespective of the underlying maturity) does not seem to be as statistically significant in Germany as it is for the U.S., as discussed in Bernanke (1990).

Let us now focus on the advantage of combining different estimated probabilities at the one- two- and three-month-ahead forecast horizons illustrated in Figure 1.4. As is clearly observable, the estimated recession probabilities of all probit specifications feature a similar pattern, though there are some periods where the range of estimated probabilities becomes particularly high. This is especially important in middle ranges of the interval [0, 1], where the signal threshold of a recession might be set. In order to assess in a more formal manner the capability of the probit regressions to deliver accurate signals for the occurrence of a recession, the percentage of Type I and Type II errors for a success cut-off value of 0.5 are summarized in Table 1.1. As these summary statistics clearly show, the accuracy in predicting especially the recessionary periods vary from correctly predicting 36 out of 69 recessionary periods (52.17 %) by EQ.B-SPRD1YG to 47 out of 69 (68.12 %) by EQ.B-SPRD5YG.

It is also interesting to note that the forecast accuracy in predicting recessions of the different probit specifications varies across the forecast horizon: At the one-month forecast horizon EQ.B-SPRD5YG has the highest forecast accuracy, at the two-month the specification with the best performance is EQ.B-SPRD3YS, and at the three-month horizon EQ.B-SPRD1YG and EQ.B-SPRD1YS deliver the best values. This indirectly highlights the value-added of combining the estimated probabilities of the alternative probit specifications since the specifications cover a certain range, which will dynamically change over time.

Out-of-Sample Evaluation

In order to assess the out-of-sample forecasting performance of the estimated probit models, following Moneta (2005) the out-of-sample recession probability forecasts were computed under real-time conditions by performing the following steps: First, the different probit regressions were estimated over the 1991:1 to 2004:5 period in order to provide a good starting estimation of the parameters. Then, the probability of recession at a given month ahead was forecasted and its value recorded. After adding one more month to the revised estimation period and dynamically re-estimating the different probit regressions, the procedure was repeated. At the end a series of out-of-sample estimated probabilities over the publications 2004:6 to 2011:8 was obtained, while for each publication the 1-, 2- and 3-month-ahead forecasts were recorded.

To evaluate the out-of-sample forecasting performance of an estimated probit model $\mathcal{M}$, three common measures of forecast accuracy (Rudebusch and Williams, 2009) were employed: the mean absolute error (MAE)

$$MAE(\mathcal{M}, h) = \frac{1}{T}\sum_{t=1}^{T} |P_{t|t-h}^{\mathcal{M}} - b_t|,$$

the root mean squared error (RMSE)

$$RMSE(\mathcal{M}, h) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(P_{t|t-h}^{\mathcal{M}} - b_t\right)^2},$$

and the Theil Inequality Coefficient

$$Theil = \frac{\sqrt{\sum_{t=1}^{T}(P^{\mathcal{M}}_{t|t-h} - b_t)^2/T}}{\sqrt{\sum_{t=1}^{T}(P^{\mathcal{M}}_{t|t-h})^2/T} + \sqrt{\sum_{t=1}^{T} b_t^2/T}}$$

which by definition lies in the interval $[0, 1]$, where 0 represents a perfect fit and 1 no explanation whatsoever.

As can be seen in Table 1.2, the estimated probability series resulting from the combination of the underlying forecasts seem to deliver not only statistically meaningful results and significant predictive power, but also feature good out-of-sample properties.[12] When comparing the two alternative combination schemes, a large congruency between the simple and the weighted average approach can be observed. In particular the timing of the recession signal is the same for almost all horizons if this signal is based on a recession probability above 50% (see Table 1.2). However, in terms of the MAE, RMSE and Theil coefficient provided in Table 1.2, the weighted average delivers better results for two of the three horizons. This can be traced back to the number of outliers in Figure 1.5, which correspond to months where the probit forecasts exceed 50% recession probability although the benchmark method does not recognize a recession in these periods (and vice versa). Both the number of these outliers and their level suggest that a policy-maker should prefer the weighted average approach. Summing up, these statistics confirm the benefits of a combination of real-time predictions with different forecast horizons, while aiming at the same forecast month. The reason for such benefits is the fact that forecasts based on the most recent, but also most uncertain information (zero revisions as the youngest considered data) can be stabilized by ones based on less recent, but also less uncertain information (first revisions as the newest considered data).

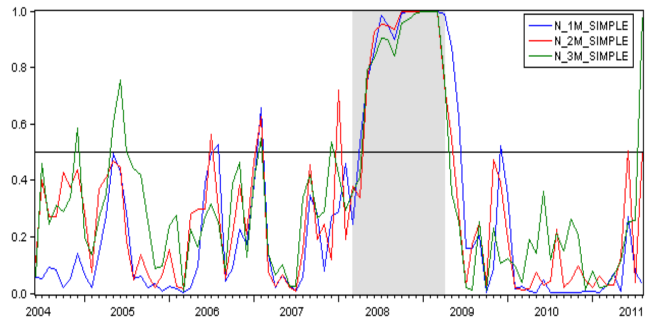
Figure 1.5 illustrates the real-time out-of-sample performance of the two combination schemes in predicting German recessions as identified by the modified Bry-Boschan algorithm (grey area, from 2008M2 to 2009M3). In this context, it is again worth noting that the real-time probability forecasts were computed before this recession dating was complete. Indeed, while this ex-post dating algorithm determines the beginning of the recession (February 2008) in September 2008, the composite probit model delivers a stable early signal for the beginning of the recession already in May 2008. This means that the essential information for a policy maker was provided four months earlier with the composite probit model than with a typical non-parametric Bry-Boschan-like instrument.

[12]Theobald (2012) shows that there is enough variation among the dynamic regression variables to justify a combinatorial approach from a statistical perspective.

Table 1.2: Statistical Evaluation Measures for Combined Real-Time Recession Probability Forecasts - GER (estimation start: 1991:1 – 2004:5, real-time out-of-sample path: 2004:6 – 2011:8)

Combination	Horizon	MAE	RMSE	Theil	Time of Signal$_{\geq 0.5, \leq 0.5}$	
simple average	1M	0.1495	0.2472	0.2946	2008M4	2009M7
	2M	0.2066	0.2881	0.3385	2008M5	2009M6
	3M	0.2384	0.3093	0.3589	2008M5	2009M5
weighted average	1M	0.1633	0.2481	0.2974	2008M2	2009M7
	2M	0.2142	0.2792	0.3322	2008M5	2009M6
	3M	0.2338	0.2965	0.3508	2008M5	2009M5

Simple Averaging - Germany



Weighted Averaging - Germany

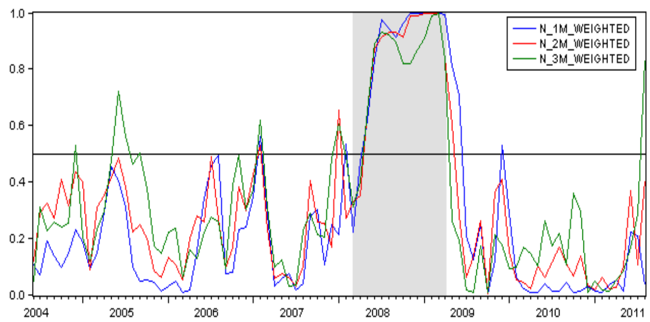


Figure 1.5: Real-time recession probabilities - Germany. The time axis is linked to the publications between 2004M06 and 2011M08, which means that the last observation of a series is given for the date of publication minus the data availability lag. The different lines represent the forecast horizons starting from the date of publication. The grey areas correspond to the ex-post recession detection by the modified Bry-Boschan (BB) algorithm. The setting for the underlying BB algorithm is a moving average with a degree equal to 2 and a severity, S_j greater than 0.05.

Additionally, the best-performing three-months-ahead forecast predicts the end of the recession for May 2009 - an impressive result when compared with the result of the ex-pos t dating provided in October 2009. Again, the composite probit model is five months ahead.

1.3.2 Real-Time Prediction of U.S. Recessions

Out-of-Sample Evaluation

In order to confirm the effectiveness of the composite probit indicator we also present real-time recession forecasts for the U.S. economy. In general, a similarly large set of financial and real economy regressors as in the German case was used. Most of the data stem from the databases of the Federal Reserve Bank of St. Louis (`www.stlouisfed.org/fred`) including its real-time database (`http://alfred.stlouisfed.org/`). The only exception is the Sentiment Index from the Conference Board (`http://www.conference-board.org/data/`). In contrast to Germany, where the structural break due to the German reunification limits the number of observations significantly, for the U.S. it is possible to estimate a longer sample comprising monthly observations from 1969:1 to 1999:12 in the in-sample period. The real-time out-of-sample path lasts from 2000:01 to 2011:08. For comparison, the official NBER recession dating is used, according to which the dotcom recession occurred between March and October 2001 (these dates being announced in November 2001 and July 2003, respectively), and the recent recession occurred between December 2007 and May 2009 (with the respective announcements in December 2008 and September 2010), as illustrated in Figure 1.6. Table 1.3 and Figure 1.7 show the real-time out-of-sample performance of the combination schemes for U.S. data.

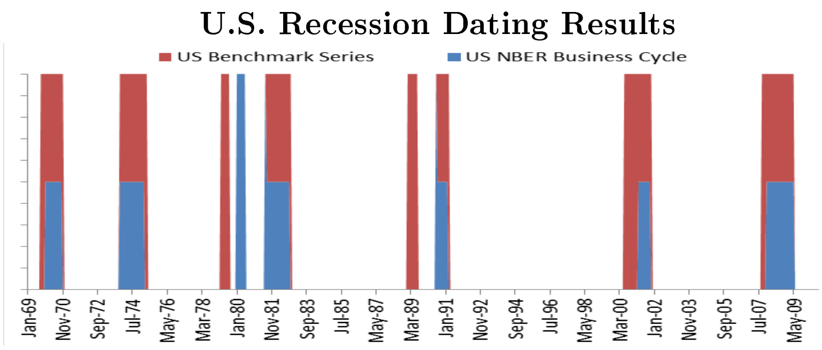
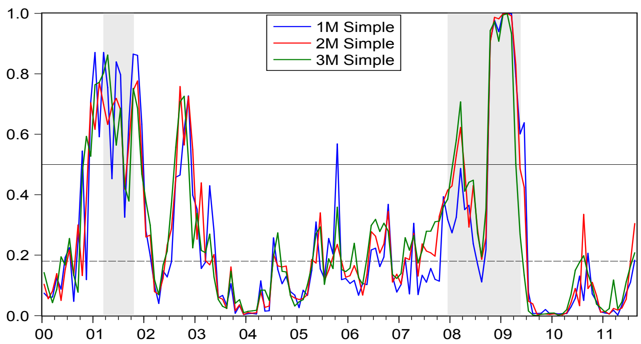


Figure 1.6: The modified BB algorithm correlates strongly with the official NBER announcements.

Table 1.3: Statistical Evaluation Measures for Combined Real-Time Recession Probability Forecasts - U.S. (estimation start: 1969:1 – 1999:12, real-time out-of-sample path: 2000:1 – 2011:8)

Combination	Horizon	MAE	RMSE	Theil	Time of Signal$_{\geq 0.5, \leq 0.5}$		Time of Signal$_{\geq 0.5, \leq 0.5}$	
	1M	0.2134	0.3155	0.3939	2000M12	2002M01	2008M3	2009M7
simple average	2M	0.2166	0.3007	0.3938	2000M12	2002M01	2008M2	2009M6
	3M	0.2264	0.3056	0.3945	2000M11	2001M12	2008M2	2009M5
	1M	0.2114	0.3043	0.3802	2000M12	2002M01	2008M3	2009M7
weighted average	2M	0.2160	0.2977	0.3818	2000M12	2002M01	2008M2	2009M6
	3M	0.2273	0.3041	0.3808	2000M11	2001M12	2008M2	2009M5

Simple Averaging - U.S.



Weighted Averaging - U.S.

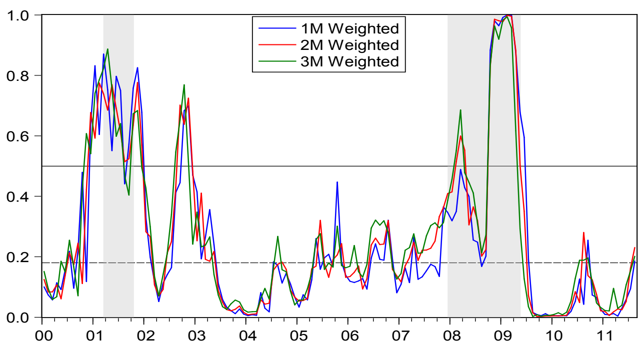


Figure 1.7: Real-time recession probabilities - U.S. The time axis is linked to the publications between 2000M01 and 2011M08. The last observation of a series is given for the date of publication minus the data availability lag. The different lines represent the forecast horizons starting from the date of publication. The grey areas represent the official NBER recessions. Early recession start and end signals are given several months before the announcements take place.

Table 1.3 reports the out-of-sample evaluation statistics for the U.S. economy. Again, the real-time prediction path resulting from the combination of the underlying forecasts features good out-of-sample properties – both in terms of the measures of forecast accuracy as well as in terms of early signals when compared to recession start and end as officially dated by the NBER.

Over the whole analyzed sample of about ten years, our composite indicator wrongly identifies four months in 2002 as recessionary[13] and six months as expansionary, if evaluated using the 50% probability level. However, these periods have been subject to substantial data revisions, and therefore a policy-maker should not concentrate on only the 50% probability level, but should also consider higher thresholds.

A particularly striking feature of Figure 1.7 is the puzzling drop of the recession probability forecasts in the middle of the last recession, where there are some months in 2008 with a recession probability below 50%. It could be argued that in most of these periods the recession probability was still above the statistical average of 19% (obtained from a probit estimation of the binary series depending only on a constant – the dotted line in Figure 1.7). However, a more economically meaningful interpretation of this finding is that the forecasting power of the yield curve (and by extension of many other financial variables) was temporarily degraded due to the direct intervention of the U.S. Federal Reserve in the bond market in the context of its quantitative easing strategy, affecting the performance of the probit models.

1.4 Comparison with Existing Approaches from the Literature

As discussed, the presented composite probit model shows convincing in-sample results consistent with previous studies such as Estrella and Mishkin (1998), Kauppi and Saikkonen (2008) and Nyberg (2010) in terms of a high congruency with the recessionary periods identified by its benchmark methodology (see in particular Figure 1.4). However, as Estrella and Mishkin (1998, p.55) rightly point out, "in-sample and out-of-sample performance can differ greatly", and this discrepancy (and in general terms, the model's performance) can vary significantly under real-time conditions (Croushore and Stark, 2001). In this section we thus explicitly compare our results, especially for the U.S., with outcomes of previous related empirical studies.

[13] A certain sensitivity towards a recession declaration is fully intended as a missed recession is treated as the statistical alpha error. It is difficult to say how far the U.S. were in 2002 from a recessionary phase which the NBER would have identified.

Table 1.4: Statistical Evaluation Measures for Existing Approaches from the Literature - US (estimation start: 1969:1 – 2007:6, real-time out-of-sample path: 2000:7 – 2011:8, forecast horizon composite dynamic probit model: 3 months, horizon all other approaches: 6 months)

Approach	MAE	RMSE	Theil	Time of Signal$_{\geq 0.5, \leq 0.5}$	
Dueker (1997), Table 1 $\varphi_t = c + \alpha SPRD10Y_{t-h} + u_t$	0.2798	0.3986	0.5851	2001M04 2001M07	2007M09 2007M10
Dueker (1997), Table 2 $\varphi_t = c + \alpha y_{t-(D_y+h)} + u_t$	0.2728	0.3861	0.5636	– – –– – – – –	2006M07 2007M06
Estrella and Mishkin (1998), Table 3 $\varphi_t = c + \alpha SPRD10Y_{t-h} + \beta SP\_DLN_{t-h} + u_t$	0.3055	0.4405	0.5636	2001M03 2001M07	2007M08 2007M09
Nyberg (2010), Equation 7 $\varphi_t = c + \alpha\varphi_{t-1} + \mathbf{x}'_{t-k}\beta + u_t$	0.1854	0.3057	0.4293	2001M04 2001M10	2007M08 2009M05
Nyberg (2010), Equation 8 $\varphi_t = c + \alpha\varphi_{t-1} + \mathbf{x}'_{t-k}\beta + y_{t-(D_y+h)}\mathbf{x}'_{t-k}\gamma + u_t$	0.2011	0.3556	0.4469	2001M04 2001M10	2007M04 2008M08
Composite Dynamic Probit Model $\varphi_t = c + \sum_{j=h+D_y}^{p} \alpha_j y_{t-j} + \sum_{j=h+D_{\mathbf{x}}}^{q} \mathbf{x}'_{t-j}\beta_j + u_t$	0.2273	0.3041	0.3808	2000M11 2001M12	2008M02 2009M05

Table 1.4, as well as Figures 1.8 and 1.9 summarize the forecasting performance of selected static and dynamic probit models proposed by Dueker (1997), Estrella and Mishkin (1998) and Nyberg (2010) (where $SPRD10Y_t$ represents the 10-year yield spread and SP_DLN_t the growth rate of stock prices) under real-time conditions. Table 1.4 reports the measures of forecast accuracy for the above-mentioned approaches; the selection of a forecast horizon of six months is based on all authors stating that their models perform "best" at this forecast horizon.[14] In particular, this implies that we compare the performance of our composite model at the three-month-ahead horizon with the performance of the other models at the six-month-ahead horizon. This of course makes a direct comparison between the different models more difficult, but since Dueker (1997), Estrella and Mishkin (1998) and Nyberg (2010) provide no automatized lag selection procedure working in real-time, the selection of another forecast horizon (such as the three-month ahead) would be as ad-hoc as the six-month ahead, and could be even interpreted as providing a negative bias with respect to forecast performance of these models as it would focus on a specification which was not considered as the "best-performing" by the original researchers.

Generally, our composite model outperforms the other models in terms of

[14] With real-time data, it is essential to understand what exactly is meant by the term forecast horizon. For instance, Nyberg (2010, p.14) discusses his specification as follows: "When the forecast horizon h lengthens, the lags of explanatory variables should be tailored so that only the information included in the information set Ω_{t-h} at forecast time $t-h$ is used. For example, when the forecast horizon is 16 months meaning that we are interested in forecasting the seven month ($h = 7$) ahead value of the recession indicator, then in the predictive models for the U.S., the vector $\mathbf{x}^{US}_{t-k}$ contains the following variables $\mathbf{x}^{US}_{t-k} = \left(SP^{US}_{t-7} SP^{GE}_{t-7} r^{US}_{t-7}\right)$."

the RSME and the Theil coefficient, and has the second best performance (after Nyberg's (2010) model) in terms of the MAE. These results are best explained by looking at the graphical forecast results of these models as depicted in Figure 1.8. The out-of-sample real-time recession probabilities forecasted by Dueker (1997) and Estrella and Mishkin (1998) are of a much lower level than those of the composite indicator (see Figure 1.7) during the two recessionary periods considered, delivering more ambiguous signals about a recession in the near future. Nonetheless, using the 50% probability level for the signaling of a prospective recession, the static approaches by Dueker (1997) and Estrella and Mishkin (1998) (which includes in addition to the yield spread the growth rate of the stock prices) are indeed able to predict the U.S. dotcom recession even under complete real-time conditions, while the dynamic specification proposed by Dueker (1997) does not. However, the composite dynamic probit model outperforms these models in terms of the absolute level of the predicted recession probabilities, as well as in terms of the persistence of the probabilities above the 0.5 threshold (see Figure 1.7) during the 2001 dotcom recession. Moreover, the composite dynamic probit model also performs better in predicting the recession in the aftermath of the financial crisis in real time, as the simpler and less flexible models deliver probabilities far below the 0.5 threshold even when the economy is already suffering significantly from recession. In our view the poor performance of these other models in comparison to our composite indicator can be traced back to a) the lack of real-economy regressors on the right-hand side and b) the lack of a flexible lag structure which in the composite dynamic model is implemented through the automatized general-to-specific and specific-to-general lag selection procedures. Indeed, by updating the lag structure for each publication, our composite model is more capable of reflecting the fact that the last recession was structurally different from its predecessors.

Kauppi and Saikkonen (2008) also analyze real-time out-of-sample predictions of both static and dynamic probit models coming to the conclusion that dynamic specifications outperform static specifications. An important finding by Kauppi and Saikkonen (2008, p.32) is that the 2001 recession was hard to predict. The recession probability forecasts of their best-performing specification only reach values above 50% around the midpoint of the recession as dated ex-post by the NBER. In contrast, our composite probit model would have been able to provide a timely recession warning, as our three-month-ahead recession probability forecast exceeds the 50% line in November 2000. Thus a timely recession prediction would have been made for February 2001 (see Table 1.3), several months before the NBER's officially declared beginning of the recession as March 2001.

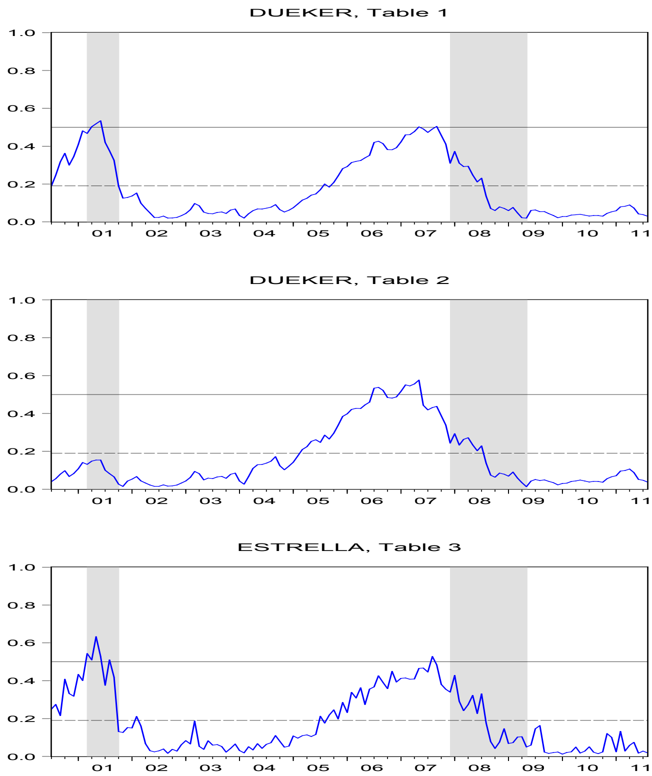


Figure 1.8: Real-time recession probabilities - U.S. The time axis is linked to the target month of the real-time forecast between 2000M01 and 2011M08, *i.e.* the date of publication plus forecast horizon (6 months). Top graph: Real-time forecasts with the static specification proposed by Dueker (1997) in Table 1: $\varphi_t = c + \alpha SPRD10Y_{t-h} + u_t$. Center graph: Real-time forecasts with the dynamic specification proposed by Dueker (1997) in Table 2: $\varphi_t = c + \alpha y_{t-(D_y+h)} + u_t$. Bottom graph: Real-time forecasts with the static specification proposed by Estrella and Mishkin (1998) in Table 3: $\varphi_t = c + \alpha SPRD10Y_{t-h} + \beta SP\_DLN_{t-h} + u_t$.

This would have meant just one month difference between our forecast and the official NBER dating, with the composite probit model even running ahead. Concerning the end of the recession, the composite probit as well as the Kauppi and Saikkonen (2008) model deliver similarly satisfactory results. We suspect the overall better prediction performance of the composite probit model to be due to the large set of leading indicators which at the initial stage are all available in the model and then for each publication are reduced to a streamlined and efficient design. In contrast, Kauppi and Saikkonen (2008) only consider the supposedly best publication-invariant predictor, i.e. the slope of the yield curve.

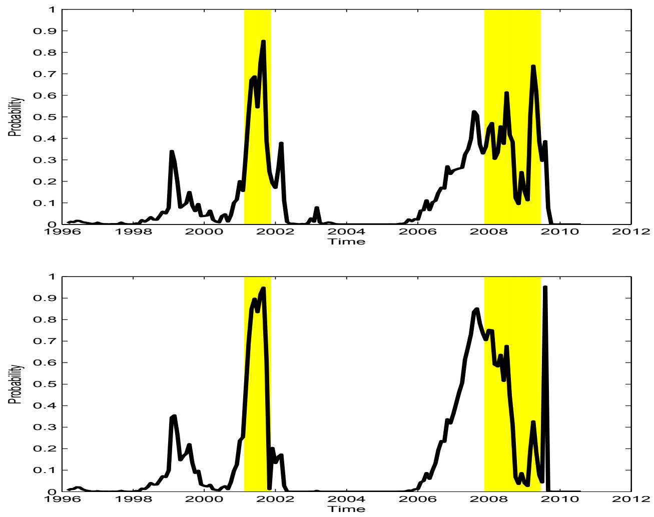


Figure 1.9: Real-time recession probabilities - U.S. The time axis is linked to the target month of the real-time forecast between 2000M01 and 2011M08, *i.e.* the date of publication plus forecast horizon (6 months). Top graph: Real-time forecasts with the autoregressive specification proposed by Nyberg (2010) in Equation 7: $\varphi_t = c + \alpha\varphi_{t-1} + \mathbf{x}'_{t-k}\beta + u_t$. Center graph: Real-time forecasts with the autoregressive interaction specification proposed by Nyberg (2010) in Equation 8: $\varphi_t = c + \sum_{j=h+D_y}^{p} \alpha_j y_{t-j} + \sum_{j=h+D_\mathbf{x}}^{q} \mathbf{x}'_{t-j}\beta_j + u_t$.

Figure 1.9 replicates real-time estimation results from Nyberg (2010). Again, it should be pointed out that the main difference between Nyberg's (2010) model and ours is the use of a larger set of regressors to reflect overall economic activity – while Nyberg (2010) mainly employs financial explanatory variables, we use pooling methods based on a combinatorial space in order to incorporate the larger set of regressors. Furthermore, while Nyberg (2010) deals comprehensively with real-time data availability lags by "adding one month to the previous estimation period and re-estimating the parameters" (Nyberg, 2010, p.14), data revisions seem to be ignored. As in the previous case, we replicate and extend the results provided by Nyberg's (2010, p.35) "best-performing" specifications, the autoregressive as well as the autoregressive interaction model, to compare them with the ones from the composite dynamic probit model.

Figure 1.9 shows that Nyberg (2010) also succeeds in producing accurate real-time out-of-sample predictions for the U.S. dotcom recession. This may emphasize the performance of financial market indicators for recession forecasting during this episode. However, when extending the real time path up to the last recession, the Nyberg specification performs worse than our composite indicator both in terms of the absolute level of the recession probability as well as in terms of misinterpreted expansionary periods in the middle of the recession. The composite probit model seems to tackle this problem more adequately by employing a larger regressor set and by taking up Timmermann's (2006) notion of averaging the inference from different specifications.

1.5 Concluding Remarks

The timely and accurate recognition of turning points in the business cycle is one of the most important, but also one of the most difficult tasks in macroeconomic forecasting. Along these lines of research, we present a practical econometric approach to forecast recession probabilities under real time conditions based on the combination of alternative dynamic probit regressions, and discuss its real-time out-of-sample performance both for the German and U.S. economies.

Although we do not provide an extensive analysis of the total space of combination schemes and thus cannot explicitly determine the overall "best" working scheme in real time, all of the considered schemes reveal a relatively high level of forecast accuracy, both in terms of measures of forecast accuracy and in terms of a graphical analysis. In particular, our composite model seems to outperform existing approaches among the same class of econometric models such as Dueker (1997), Estrella and Mishkin (1998) and Nyberg (2010). This is due to a) the use of both real-economy and financial market regressors as explanatory variables, b) the flexible lag structure which results from the automatized general-to-specific and specific-to-general lag selection procedures, and c) the combination of various forecasting models.

Finally, several extensions of the composite probit model can be elaborated. Dueker (1997) for instance proposed a coefficient variation scheme via Markov Switching in the probit model. But, while he obtains reasonable results for this specification, it is not clear whether these results can be easily reproduced under complete real-time conditions, i.e. when explicitly considering publication lags, data revisions and a potentially limited real-time data history such as in the German case because of reunification. Nonetheless, accounting for potential asymmetries in the predictive power of leading indicators, as done e.g. by Nyberg (2010), in an automatized framework such as ours, seems a promising line of research to pursue.

Appendix A

A.1 Real-time Forecasts with Lagged Values of the Business Cycle Reference Series and of the Binary Recession Series

Table 1.5: Statistical Evaluation Measures for Combined Real-Time Recession Probability Forecasts - GER (estimation start: 1991:1 – 2007:8, real-time out-of-sample path: 2007:9 – 2011:8)

Combination	Horizon	MAE	RMSE	Theil	Time of Signal$_{\geq 0.5, \leq 0.5}$	
simple average	1M	0.1302	0.2315	0.2036	2008M4	2009M5
	2M	0.1363	0.2469	0.2158	2008M5	2009M5
	3M	0.1296	0.2429	0.2163	2008M5	2009M4
weighted average	1M	0.1429	0.2496	0.2193	2008M4	2009M5
	2M	0.1296	0.2346	0.2080	2008M5	2009M4
	3M	0.1289	0.2223	0.2028	2008M5	2009M3

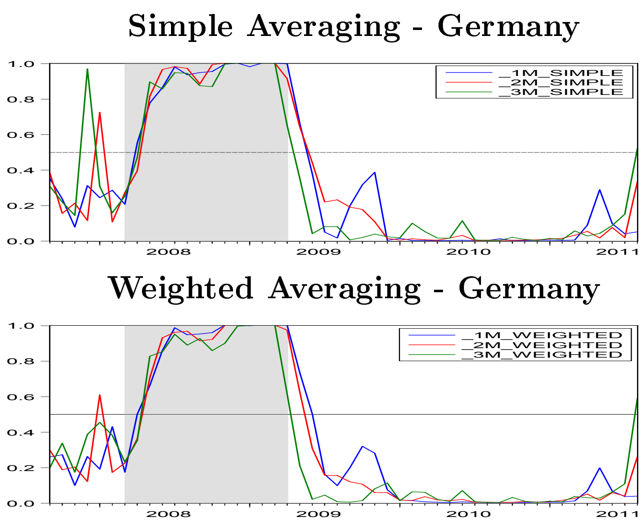


Figure 1.10: Real-time recession probabilities - Germany. The regression equation includes both the lagged binary as well as the lagged reference series. The setting for the underlying modified BB algorithm is a moving average degree of 2 and a severity, S_j greater than 0.025.

A.2 Real-time Recession Forecasts Based on a Higher Degree of Smoothing of the Business Cycle Reference Series

Table 1.6: Statistical Evaluation Measures for Combined Real-Time Recession Probability Forecasts - GER (estimation start: 1991:1 – 2007:8, real-time out-of-sample path: 2007:9 – 2011:8)

Combination	Horizon	MAE	RMSE	Theil	Time of Signal$_{\geq 0.5, \leq 0.5}$	
simple average	1M	0.1979	0.2989	0.2973	2008M4	2009M4
	2M	0.2078	0.3261	0.3115	2008M5	2009M4
	3M	0.2661	0.3676	0.3549	2008M4	2009M4
weighted average	1M	0.1854	0.2745	0.2750	2008M4	2009M4
	2M	0.2119	0.3128	0.3062	2008M5	2009M4
	3M	0.2672	0.3637	0.3615	2008M4	2009M4

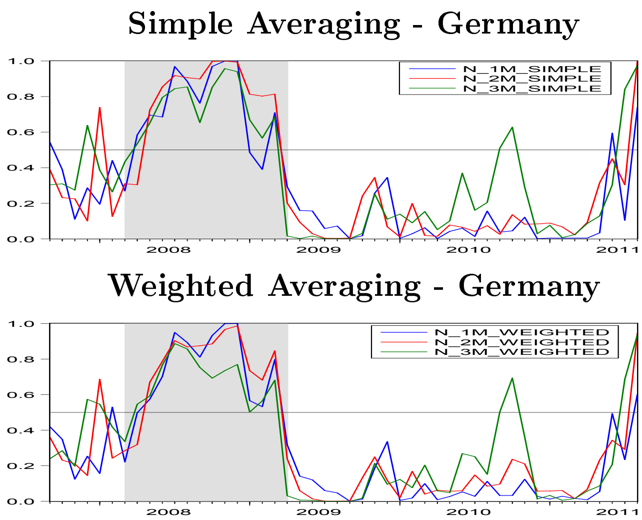


Figure 1.11: Real-time recession probabilities - Germany. The time axis is linked to the publications between 2007M09 and 2011M08, which means that the last observation of a series is given for the date of publication minus the data availability lag. Here, the regression equation includes both the lagged binary as well as the lagged reference series. The setting for the underlying modified BB algorithm is a moving average degree of 3 and a severity, S_j greater than 0.05.

A.3 Recession Prediction using the Foreign Yield Spread

We discuss in the following the question as to whether other explanatory variables could be included based on Nyberg (2010), who, in particular for Germany, finds that the foreign term spread with the U.S. has some predictive power. Figure 1.12 shows the explanatory contributions *excon* of the foreign term spread. The contributions are linked to the fitted value of φ which can easily be computed for a single specification: it is the inverse of the cumulative normal density function $\Phi^{-1}(.)$[15]. The values illustrated in Figure 1.12 are the average of different specifications for a certain forecast horizon, while the time axis displays different publications.[16]

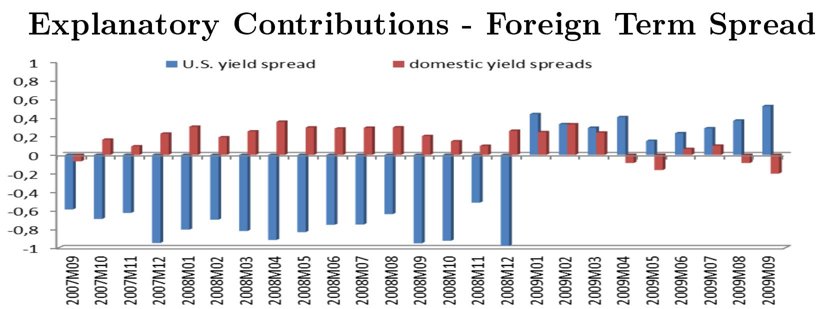


Figure 1.12: Averaged explanatory contributions of selected variables related to the fitted values of φ, i.e. for $\mathbf{x}_t = \{\text{significant lags of a certain regressor}\} : excon = \mathbf{x}_t'\beta / |\varphi_t|$. Normalized scaling method. The time axis is linked to the publications between 2007M09 and 2009M09. The quintessence is that no gaps arise reflecting the fact that for each and every publication certain lags, at least in one of the underlying specifications, turn out to be highly significant, i.e. $p < 0.05$.

As illustrated in Figure 1.12, both the U.S. and the domestic term spreads are persistently significant, while simultaneously revealing high-level explanatory contributions confirming the results by Bernand and Gerlach (1998) and Nyberg (2010). However, the inclusion of both series in the same set of regressors is also likely to create some multicollinearity problems. For instance, Figure 1.12 is based on estimations in which the three-month EURIBOR had to be dropped to include the U.S. spread. See also Estrella and Mishkin (1998, p.55) who discuss the problem of overfitting.

[15] A fitted value equal to 0 corresponds to a recession probability of 50% and so on. In contrast, in Figure 1.12 a negative sign corresponds to a recession contribution, i.e. $\Phi(excon \times |\varphi|) > 50\%$.

[16] Note that applying $\Phi(.)$ to the average of the $\varphi^i, i = 1 \ldots 20$ (and implicitly this also concerns the explanatory contributions) is not equal to the (weighted) average of the recession probabilities, where the latter one represents our combinatorial approach. Nevertheless this way of presentation provides some useful insights.

Chapter 2

Markov Switching with Endogenous Number of Regimes and Leading Indicators in a Real-Time Business Cycle Application

2.1 Introduction

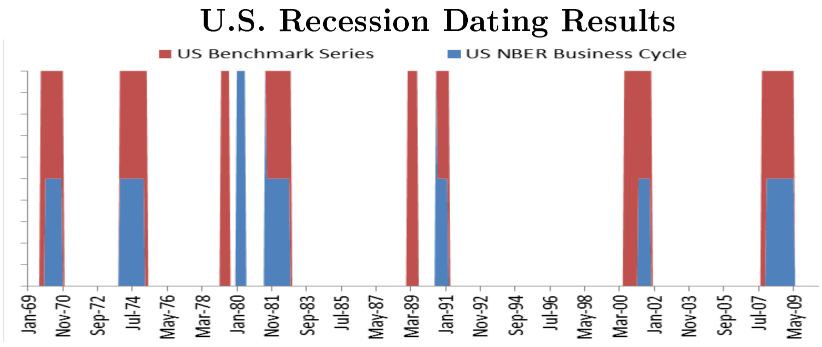
As a consequence of the financial crisis in the years 2008 and early 2009, Germany suffered its strongest decline in GDP since the global economic crisis of 1929. In such a situation the efficiency of economic policies strongly depends on a timely detection of the recession and on insights about its potential deepness as early as these are available. Related to these challenges, this chapter delivers real-time predictions within a Markov Switching (MS) framework. In the class of regime switching models it is also possible to analyze the impact of exogenous regressors by the Markov Switching Autoregressive Models with Exogenous Variables (MSARX). Among others Lee, Liang and Chou (2009) use such an approach for regressing a proxy variable of the real estate cycle on its lags as well as on a composite leading index. Still, most of the MS business cycle literature concentrates on purely autoregressive estimations following the famous Markov Switching Mean Model (MSM) by Hamilton (1989), or as stated by Boldin (1996, p.1): 'Because the estimated parameters of relatively simple MSM specifications match many stylized facts about the business cycle, this framework has become an important alternative to linear, autoregressive structures.' Contrary to the linear case, a straightforward set of specification tests for MS models - in particular covering a highly parameterized design - with results clearly supporting the MSM, are not available. Although some test procedures have been developed, see Breunig, Najarian and Pagan (2003) and Carrasco, Liang and Plogerer (2013), it is difficult to align them with a more parameterized design. But in contrast to a purely autoregressive MSM, the inclusion of leading indicators as explanatory variables can be appealing due to promising additional information for policy makers.

The chapter at hand considers this additional information in bivariate MSARX regressions with lagged dependent and (lagged) leading indicator variables. The forecast of business cycle turning points represents the central feature of the model. With respect to this aim, different MSARX specifications allow us to take advantage of averaging effects (pooling methods) and as an additional feature to endogenize the number of model-inherent regimes depending on available real-time data. Indeed, it turns out that both of our extensions, the combination of different MSARX specifications as well as the real-time change in the number of regimes can significantly improve the levels of forecast accuracy.

The chapter is structured as follows: Section 2.2 provides a short overview of the literature related to real-time business cycle forecasts. Section 2.3 reviews elements of the filter introduced by Hamilton (1989, 1990) in order to stress the fact that exogenous variables have a numerical impact on the state probabilities when using an expectation-maximization (EM) algorithm. Section 2.4 describes data input and in-sample results of the model. Also in this section, model specification is discussed in detail, illustrating how the structure of the model results from a compromise between information needs and available real-time data records. While the former suggest a super highly parameterized design, the latter clearly restricts the parameter space to some extent. Section 2.5 deals with the out-of-sample results. First we introduce the notion of how to change the number of regimes in real-time. Next, we present real-time out-of-sample results for the industrial production as the monthly dependent variable. Finally the procedure is repeated for German data of the OECD Composite Index of Leading Indicators (CLI). Section 2.6 concludes.

2.2 The Literature

Back in the 1970s, research efforts were mainly focused on exact dating of business cycle turning points represented by the seminal book by Bry and Boschan (1971), where they developed a solid working non-parametric dating algorithm. The output of a modified version of this algorithm, including the



The modified BB algorithm correlates strongly with the official NBER announcements. Source: Proaño and Theobald (2012).

Harding and Pagan (2002) triangle approximation approach, will serve as a benchmark for our MSARX forecast results. The algorithm works as an ex-post-dating procedure which is able to deliver reliable recession signals after the recession recognition lag (5 months) and the publication lag (2 months) of the reference series (industrial production) have expired. For the U.S. the so-dated recessions overall coincide with the official NBER announcements, see Figure **??**. For Germany, where no official business cycle chronology exists, this procedure fills the functional gap of transferring the reference series into a binary series of business cycle phases (1 for a recessionary phase, 0 for a non-recessionary phase).

Nowadays the focus has turned to real-time business cycle predictions ahead of the publication point in time. Therefore it is crucial to deal with two questions: Firstly, what estimation procedure to use, and secondly which indicators related to the business cycle to be included. *Inter alia* the development of estimation procedures was fostered by Chauvet and Potter (2005) using different specifications of the probit model, by Stock and Watson (1989) introducing dynamic factor models for the business cycle, and by Hamilton (1989) proposing the Markov Switching model which this chapter extends to incorporate a combination of different MSARX specifications. The resulting model is then applied to monthly German real-time data. Simultaneously, within the development of different prediction procedures the set of leading real economy indicators was extended to include financial ones such as spreads from the term structure of interest rates, e.g. by Estrella and Hardouvelis (1991) and the spread between corporate and public issuers, e.g. by Friedman and Kuttner (1992). Most recently the connection between the corporate spread and economic development has been analyzed by Gilchrist, Yankov and Zakrajsek (2009). Disparity between the characteristics of financial and real economy indicators becomes particularly essential with real-time forecasts. While financial data, at least with monthly frequency, is provided immediately and is not subject to revisions, this is not the case for most of the real economy variables as they are subject to a publication lag and data revisions. As Diebold and Rudebusch (1991) pointed out for the U.S. Composite Leading Index, revisions and the lagged data availability substantially affect the predictive power of leading indicators. That is why this chapter considers real-time data and additionally contributes to the literature by adjusting the MS model while proceeding on the real-time path. Whereas pooling methods in the spirit of Timmermann (2006) are regularly employed with other real-time forecast procedures, for instance see Proaño and Theobald (2012) for a combinatorial approach of different dynamic probit models, to the best of our knowledge no Markov Switching literature follows such an approach. Indeed it is the basis for showing how a data sample-dependent change in the number of regimes can significantly improve real-time forecasts.

2.3 The Empirical Model

When starting business cycle modeling, it is useful to look for a well-defined and generally acknowledged borderline between recessions and expansions such as the one given by the National Bureau of Economic Research (2011) for the U.S.:

> A recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales. A recession begins just after the economy reaches a peak of activity and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion.

At first glance this definition may suggest a five-dimensional MSVAR model, as introduced by Krolzig (1997). But MSVAR specifications would additionally enlarge the parameter space beyond all the extensions we are introducing in this chapter. Given the currently available German real-time data records such an approach is likely to fail. For the same reason one-dimensional equations are arranged to include only two kinds of regressors, lags of the dependent variable and one exogenous variable (leading indicator) including its lags. Moreover, only the coefficient of the most recent lag is chosen to switch in order to minimize the number of parameters that have to be estimated. This leads to the following form

$$\begin{aligned} y_{t+h} &= \beta_0^{S(t)} + \beta_{1,y}^{S(t)} y_{t-D_y} + \sum_{j=2}^{p} \beta_{j,y} y_{t-(j-1)-D_y} \\ &\quad + \beta_{1,x}^{S(t)} x_{t-D_x}^i + \sum_{j=2}^{q} \beta_{j,x} x_{t-(j-1)-D_x}^i + u_t, \\ u_t &\sim N\left(0, \sigma^S(t)\right), t = 1, \ldots, T, i = 1, \ldots, I, \\ z_t^{'} &= \left(1, y_{t-D_y}, \ldots, y_{t-p+1-D_y}, x_{t-D_x}^i, \ldots, x_{t-q+1-D_x}^i\right), \\ \beta^{S(t)'} &= \left(\beta_0^{S(t)}, \beta_{1,y}^{S(t)}, \beta_{2,y}, \ldots, \beta_{p,y}, \beta_{1,x}^{S(t)}, \beta_{2,x}, \ldots, \beta_{q,x}\right), \\ \theta^{'} &= \left(\beta^{S(t)'}, \sigma^{S(t)}, (p_{kl})\right), \end{aligned} \tag{2.1}$$

where h represents the forecasting horizon and D_y, D_x the data availability lag of the dependent and independent variable. S_t stands for the latent states that generate the total process of the observed reference series y_t. Central for the Markov Switching model is that the hidden states of the dependent variable are generated by a first order Markov chain, whose transition matrix for a two regime setting will look like

$$P = (\rho_{kl}) = \begin{pmatrix} \rho\left(S(t) = 1|S(t-1) = 1\right) & \rho\left(S(t) = 1|S(t-1) = 2\right) \\ \rho\left(S(t) = 2|S(t-1) = 1\right) & \rho\left(S(t) = 2|S(t-1) = 2\right) \end{pmatrix}, \tag{2.2}$$

where column l stands for the business cycle regime from which to jump and row k for the business cycle regime in which to jump. Later on, the model will be extended to four regimes and different intensities of the business cycle regime, so that for the transition matrix alone 12 parameters have to be estimated. This is where the available real-time data records become especially relevant. For reasons of simplicity let us stick to the two regime case where the second element in identifying the Markov chain is the starting distribution, *i.e.*

$$\xi_{1|0} = \begin{pmatrix} \rho\left(S(1) = 1|y_0\right) \\ \rho\left(S(1) = 2|y_0\right) \end{pmatrix}. \tag{2.3}$$

y_0 cannot be observed, so that one needs an initial guess for the starting distribution. In fact at the beginning of the maximization algorithm, the entries for both the transition matrix and starting distribution are chosen uniformly, which becomes important when deciding which regimes are related to recessions and which to expansions of the business cycle.

Applying the proposition of total probability delivers preliminary values for smoothed state probabilities $\xi_{t|T}, t = 1, \ldots, T$ as they were introduced by Kim (1994). Maximization with respect to θ leads the process into its second loop and the whole procedure iterates up to the convergence of the likelihood function. Concerning the estimation methodology, we want to stress the fact that with different exogenous explanatory variables, $x^i, i = 1, \ldots, I$, the solution for θ also numerically changes. Thus it is a double effect of different entries in $z_t^{'}$ and in $\beta^{S(t)}$ during the whole maximization process that ultimately generates different state probabilities for the different MSARX specifications. This effect can be directly traced back to the normal density vector

$$\begin{pmatrix} f\left(y_{t+h}|S\left(t\right)=1, z_t, \theta\right) = \frac{1}{\sqrt{2\pi\sigma^1}} \exp\left(-\left(\frac{y_t - z_t^{'}\beta^1}{2\sigma^1}\right)^2\right) \\ f\left(y_{t+h}|S\left(t\right)=2, z_t, \theta\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\left(\frac{y_t - z_t^{'}\beta^2}{2\sigma^2}\right)^2\right) \end{pmatrix}, \tag{2.4}$$

used during the maximization process. The whole model is constructed in the spirit of Timmermann (2006) assuming that each regression equation (or in other words specification) $i = 1, \ldots, I$ may be subject to a 'misspecification bias of unknown form' and that averaging can lower the effect of such a bias and hence applying pooling methods will increase forecast accuracy. As already mentioned, one reason for such a bias could be the limit of the parameter space given the available real-time data records. In the following, each equation will be represented by its exogenous variable. Against the background of trying to achieve proper forecasts, this variable is a carefully chosen leading indicator of the business cycle.

2.4 Specification and In-Sample Evaluation

2.4.1 The Data

When it comes to the question of frequency, the chapter at hand selects monthly data in order to be ready for forecasts between the quarterly publications of GDP in Germany. Fritsche and Stephan (2002) point out that the highest-correlated monthly proxy of overall German economic activity is industrial production, which captures the volatile parts of GDP such as investments and exports. Hence we select industrial production [1] to serve as the dependent variable in the equations of Section 2.3. In order to evaluate the combined forecast later on, a non-parametric ex-post-dating algorithm based on the work of Bry and Boschan (1971) and Harding and Pagan (2002) will be employed. In the tradition of this kind of business cycle literature, industrial production is also referred to as the reference series. All financial and real economy data stem from Deutsche Bundesbank (2010*a*,*b*,*c*). As a survey variable we use data provided by ifo Institut für Wirtschaftsforschung (2011) and we take OECD (2011) as the relevant data source for a German composite leading index.

[1] Given the high volatility of monthly data and as usual in related studies, we select a slightly backwards smoothed version of industrial production. For details see Theobald (2013).

BDS Test for growth rates of industrial production

Dimension	BDS Statistic	Std. Error	z-Statistic	Normal Prob.	Bootstrap Prob
2	0.051300	0.007325	7.003095	0.0000	0.0000
3	0.082194	0.011679	7.037683	0.0000	0.0000
4	0.097266	0.013956	6.969353	0.0000	0.0000
5	0.095695	0.014599	6.554701	0.0000	0.0000

Raw epsilon	0.007572		
Pairs within epsilon	25952.00	V-Statistic	0.703993
Triples within epsilon	3866120.	V-Statistic	0.546225

Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	9845.000	0.542574	12718.00	0.700909	0.491274
3	7616.000	0.424172	12556.00	0.699304	0.341978
4	6100.000	0.343352	12513.00	0.704323	0.246086
5	4885.000	0.277904	12505.00	0.711401	0.182210

Figure 2.1: BDS Test for growth rates of smoothed industrial production. The test results in a clear rejection of the i.i.d - hypothesis and confirms the appropriateness of a non-linear model. When computing ϵ/σ this lies in the interval of $[0.5, 2]$. This, as well as a maximum embedded dimension of $m = 5$, represent the relevant range for i.i.d - series according to Brock et al. (1996). Since N/m is not large, additional bootstrapped p-values are calculated.

In order to avoid the structural break linked to the German reunification, all data starts in the early 90's. At the time of writing, the corresponding publication lag of industrial production spans two months.

As comprehensively discussed by Krolzig (1997, p.20) MS in general (and in particular if highly parameterized as in this chapter) represents a non-linear model as it is, apart from special cases, not possible to find a $MA\infty$-representation. Thus, before starting to estimate, it is reasonable to apply a nonlinearity test to the reference series. A widely used test for nonlinearity is the one developed by Brock et al. (1996) which also Maheu and McCurdy (2000) apply to test for the appropriateness of their MS framework. In general the BDS tests for independence analyses whether there is any non-linear dependence in the time series after a linear ARMA model has been fitted. In detail, it is analyzed whether the residuals could follow an i.i.d process under the null (H_0: The residual time series is i.i.d.). As Figure 2.1 shows, H_0 is rejected on every regular level of significance, which in fact suggests the need to apply a non-linear method to the data.

	Sample: 1994:03 – 2010:09	Publication lag	Revision	ADF test result
Y	Industrial Production	2 months	yes	I(1)
X	Foreign Orders	2 months	yes	I(1)
X	Domestic Orders	2 months	yes	I(1)
X	Construction Permits	2 months	no	trend-stationary
X	CDAX	0 month	no	I(1)
X	Corporate Spread	0 month	no	I(1)
X	Euribor - 3M	0 month	no	I(1)
X	ifo Business Climate	0 month	no	I(0)
X	Credit Growth	1 month	no	I(1)
X	Maturity Spread	0 month	no	I(1)
X	Job Vacancies	0 month	no	I(1)

Table 2.1: Real-time characteristics and stationarity properties of selected variables.

Having dealt with the reference series, we now turn to the leading indicator series. We selected foreign and domestic orders, construction permits, CDAX stock index, the spread between corporate and public issuers' current yield, the 3-month EURIBOR interest rate, the ifo business climate index, credit growth, the maturity spread between 10-year federal bonds and 3-month EURIBOR as well as job vacancies. Hence, I introduced in Section 2.3 equals 10. Apart from job vacancies and the corporate spread, which was already mentioned in Section 2.2 because of its capability for early signaling, a similar information set contributes to the U.S. Composite Index of Leading Indicators.

The lead of most of the selected indicators is obvious since they reflect pre-stages to the production process, such as the orders, or expectations of economic development, such as the business climate. As it turns out later on, job vacancy coefficients are only weakly significant. Nevertheless its purpose is also to include at least one variable from the labour market. The reasoning behind the corporate spread is that whenever a recession approaches, this will lead to higher default rates of companies, whereas federal bonds remain a safe haven. Since short-term interest rates react more sensitively to the current economic situation, the spread between long-term and short-term maturity implies predictive potential - sometimes even ending up with an inverse yield curve. All data is calendar and seasonally adjusted. In general, financial and survey variables are not subject to revisions and to lagged data availability, whereas real economy variables are. For the real economy variables used here, the publication lag is taken to be two months - except for job vacancies, which are provided immediately. Finally, Table 2.1 displays Augmented Dickey Fuller (ADF) test results which suggest that one should difference most of the variables including the reference series. Similar non-stationarity properties have been found by Levanon (2010). For some of the series such as the term spread, there is a controversial literature whether to difference or not. In these cases we decided to potentially over-difference which is usually appraised as a feasible approach for prediction models.

2.4.2 Model Specification

The usual trade-off between improving the overall fit by additional significant regressors and making it worse by over-specification arises particularly with MS models. Whereas in a standard linear estimation an equation with only two kinds of regressors and restricted lag selection may lead to an omitted variable bias, in a combination of MSARX specifications this is only obvious when the bias occurs in each of the regime dependent equations. However, extending the equations with additional variables or switching parameters in such a way that the optimization process described in Section 2.3 only finds local maxima seems to be the greater danger, see Boldin (1996). Therefore, as it becomes obvious by means of the degree of freedom in Table 2.3, we accept at most four switching coefficients in a single equation in order to guarantee the numerical robustness of the approach. Nevertheless, in general restricted real-time data records still remains a problem. In the following we explain which coefficients we chose to switch and why. The first MS business cycle model introduced by Hamilton (1989),

$$y_t - \beta^{S(t)} = \sum_{j=1}^{4} \phi_j \left(y_{t-j} - \beta^{S(t-j)} \right) + \epsilon_t, \quad \epsilon_t \sim N\left(0, \sigma^2\right), \quad S(t) = 1, 2 \tag{2.5}$$

is helpful in deciding whether the intercept in each equation of Section 2.3 should switch or not. In particular, Perlin (2012) showed that estimating the following version of the Hamilton model

$$y_t = \beta^{S(t)} + \epsilon_{a,t}, \quad \epsilon_{a,t} \; N\left(0, \sigma_a^2\right), \quad S(t) = 1, 2 \tag{2.6}$$

$$\epsilon_{a,t} = \sum_{j=1}^{4} \phi_j \epsilon_{a,t-j} + \epsilon_{b,t}, \quad \epsilon_{b,t} \; N\left(0, \sigma_b^2\right) \tag{2.7}$$

delivers similar values for the state probabilities and switching coefficients. Indeed, this two-stage estimation identifies the switching intercept as the most relevant part for the iteration of the state probabilities and explains the inclusion of a switching intercept in our model.

In addition, we include more switching coefficients in each of the equations. Reasons for that have already been touched upon in the introduction. Technically, they are as follows: Firstly, a switching variance of the error term allows for applying the Welch test to identify different normal distributions when turning to a higher regime setting, see Section 2.5.1. Hence, the selected design implicitly delivers a way of testing the change in the number of regimes. Secondly, at least one switching coefficient of the embedded leading indicator (and of the lagged dependent variable for consistency) allows for measuring how much the coefficients change between the regimes representing different intensities of the same kind of business cycle phase (weak and strong recession or weak and strong expansion). Table 2.2 illustrates the final MSARX specifications according to a general-to-specific lag choice by information criteria. Lags can be selected up to a maximum of 5 when in addition a minimum of state probabilities agree with the available time-consistent benchmark series generated by the non-parametric algorithm. In a former version, lag choice was implemented to be renewed for each publication on the real-time path. But because of an exploding running time the lag structure in Table 2.2 was fixed for all real-time estimations.[2]

2.4.3 In-Sample Evaluation

Tables 2.2 - 2.4 and Figure 2.2 present in-sample results for the MSARX regressions. Table 2.2 already deals with a four regime setting and provides standard errors for each of the coefficients in parentheses. Standard errors have been calculated in line with the implementations by Perlin (2012) to be robust to heteroskedasticity and autocorrelation according to the methodology introduced by Newey and West (1987). Most of the coefficients are highly significant. The only instance of systematic insignificance is the rejection of the third regime's parameters when running the specification with job vacancies. In this case it cannot be excluded that each of the regime's coefficients and with this the contribution to the fit could be equal to 0. With respect to the consistency of the whole system and to the intention to include at least one variable from the labour market these results were tolerated.

Concerning the in-sample results of real-time estimations, it must be mentioned that statements about significance are linked to a certain publication - in this case to the publication on November 2010. Indeed p-values may change (slightly) for each and every publication, but it goes beyond the scope of this chapter to consider each of the publications for the in-sample analysis. Figure 2.2 illustrates the state probabilities iterated out for each of the MSARX specifications. Indeed they do change with different exogenous explanatory variables, as was theoretically anticipated in Section 2.3.

[2]First of all, lag choice is based on the publication from January 2007, *i.e.* the last before real-time forecasts of Section 2.5.2 start, but at this point in time a higher number of regimes is only hypothetical and not a result attributable to the criterion (2.9). That is why in-sample results presented in Section 2.4.3 help to check if the selected lag structure will still work for a publication (November 2010) after the change in the number of regimes has taken place.

Sample: 1994:03 – 2010:09	Switching Intercept				Switching Endogenous Lag			
Publication: 2010:11	r1	r2	r3	r4	r1	r2	r3	r4
Purely Autoregressive	0.010***	0.004***	-0.002**	-0.040***	-	-	-	-
	(0.0018)	(0.0011)	(0.0011)	(0.0008)	(-)	(-)	(-)	(-)
Foreign Orders	0.005***	0.000	-0.020***	-0.034***	-0.092***	-0.132***	-1.281***	0.349***
	(0.0003)	(0.0003)	(0.0002)	(0.0000)	(0.0000)	(0.0021)	(0.0141)	(0.0000)
Domestic Orders	0.013***	0.003***	-0.002***	-0.033***	-	-	-	-
	(0.0001)	(0.0002)	(0.0006)	(0.0009)	(-)	(-)	(-)	(-)
Construction Permits [8]	0.005***	0.005***	-0.002***	-0.039***	-0.100***	0.455***	-0.121***	-0.069***
	(0.0005)	(0.0009)	(0.0005)	(0.0006)	(0.0146)	(0.0101)	(0.0173)	(0.0156)
CDAX	0.006***	-0.001**	-0.010***	-0.031***	-0.101***	-0.094***	0.220***	0.977***
	(0.0003)	(0.0003)	(0.0001)	(0.0008)	(0.0019)	(0.0026)	(0.0018)	(0.0022)
Corporate Spread	0.010***	0.003***	-0.003	-0.054***	-0.057***	-0.089***	-0.122***	-0.420***
	(0.0002)	(0.0007)	(0.0022)	(0.0031)	(0.0068)	(0.0158)	(0.0002)	(0.0700)
Euribor - 3M	0.012***	0.004***	-0.003***	-0.043***	-	-	-	-
	(0.0032)	(0.0058)	(0.0017)	(0.0010)	(-)	(-)	(-)	(-)
ifo Business Climate	0.010***	0.009***	0.006***	0.000	0.274***	1.087***	-0.085**	-0.127***
	(0.0008)	(0.0001)	(0.0003)	(0.0007)	(0.0431)	(0.0677)	(0.0411)	(0.0264)
Credit Growth [9]	0.010***	0,005	-0.003	-0.043***	-0.117*	-0.122***	-0.128**	-0.232
	(0.0014)	(0.0264)	(0.0021)	(0.0071)	(0.0663)	(0.0273)	(0.0648)	(0.2028)
Maturity Spread	0.010***	0.003***	-0.002***	-0.027***	-	-	-	-
	(0.0010)	(0.0006)	(0.0005)	(0.0043)	(-)	(-)	(-)	(-)
Job Vacancies	0.012**	0.003**	-0.002	-0.016***	-0.123**	-0.095**	-0.175	-0.306***
	(0.0046)	(0.0013)	(0.0065)	(0.0024)	(0.0481)	(0.0417)	(0.1692)	(0.0233)

Table 2.2: Coefficients and significance in bivariate MSARX regressions with 4 regimes (r1 - r4) for the publication 2010:11. The table deals with the switching intercept and autoregressive term. * stands for $p < 0.10$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are in parentheses.

Sample: 1994:03 – 2010:09	Switching Exogenous Lag				Additional	Additional		Degree of
Publication: 2010:11	r1	r2	r3	r4	Endo. Lag	Exogenous Lags		Freedom
Purely Autoregressive	-	-	-	-	-0.113***	-	-	162
	(-)	(-)	(-)	(-)	(0.0090)	(-)	(-)	
Foreign Orders	0.015**	0.045***	0.737***	-0.078***	-	0.032***	-	154
	(0.0051)	(0.0057)	(0.0050)	(0.0000)	(-)	(0.0000)	(-)	
Domestic Orders	0.159***	0.052***	0.050**	0.053***	-	0.036**	-	158
	(0.0072)	(0.0073)	(0.0242)	(0.0053)	(-)	(0.0141)	(-)	
Construction Permits [8]	-0.003	0.159***	-0.0028*	-0.209***	-0.017	-0.002	-0.004*	148
	(0.0034)	(0.0159)	(0.0017)	(0.0033)	(0.0126)	(0.0031)	(0.0022)	
CDAX	-0.007**	-0.005***	0.136***	0.289***	-	-0.005***	-	154
	(0.0023)	(0.0003)	(0.0009)	(0.0024)	(-)	(0.0011)	(-)	
Corporate Spread	0.041***	0.018**	-0.135***	0.197***	-	-	-	155
	(0.0091)	(0.0089)	(0.0109)	(0.0262)	(-)	(-)	(-)	
Euribor - 3M	0.654***	0.043***	-0.102***	1.389***	-	-	-	159
	(0.0497)	(0.0103)	(0.0282)	(0.0279)	(-)	(-)	(-)	
ifo Business Climate	1.569***	0.206***	0.088**	0.103***	-	-	-	155
	(0.0506)	(0.0546)	(0.0378)	(0.0216)	(-)	(-)	(-)	
Credit Growth [9]	0.025**	0.009*	0.012	0.059	-	0.001	-0.010***	93
	(0.0253)	(0.0046)	(0.0111)	(0.0617)	(-)	0.0029	0.0028	
Maturity Spread	-0.043	0.018	0.006	-0.204***	-	-	-	159
	(0.0310)	(0.0130)	(0.0052)	(0.0537)	(-)	(-)	(-)	
Job Vacancies	0.068***	0.020	0.040	0.862***	-	0.048**	0.019	153
	(0.0210)	(0.0156)	(0.1057)	(0.0821)	(-)	(0.0217)	(0.1211)	

Table 2.3: Coefficients and significance in bivariate MSARX regressions with 4 regimes (r1 - r4) for the publication 2010:11. The table deals with the remaining parts of the regression equations in particular those coefficients of the exogenous variables. * stands for $p < 0.10$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are in parentheses.

Moreover, it is reasonable to assume that the predictive power of a single indicator, and thus of the corresponding specification, varies in different recessions so that the ex-ante best forecast need not be the same for the next publication. In a nutshell, the changing regime probabilities with different MSARX specifications make it advisable to rely on a combined forecast.

Each MSARX specification identifies the recession linked to the financial crisis as an additional regime, see Figure 2.2. Thus, when turning to the out-of-sample forecasts, the task will be to allow for a flexible number of regimes. But how can different regimes, in particular in the case of more than two, be identified as recessionary or expansionary states? In general, it makes sense to relate regimes to recessions or expansions. The latter determine the business cycle and with the MS model the iterated regimes generate the reference series which again is a proxy for the business cycle. Under certain conditions the identification can focus on the switching intercept. Each iteration starts with uniformly distributed initial values, so it is not necessary that the same regime label, e.g. r1, always represents the same business cycle state for each and every publication. Hence, in a two regime setting, the higher intercept will identify the expansionary regime and the lower the recessionary. In Section 2.5.1 we will discuss the more sophisticated four regime case, where different intensities of expansion and recession come into play.

Another interesting question arising with different MSARX specifications is whether a leading indicator can really bequeath its predictive power to the corresponding specification. Hints of such a relation can be found in Figure 2.2. For instance the presumed continuous recession between February 2001 and September 2003 in the specifications of CDAX and the ifo Business Climate reflects the course of the corresponding indicators. Another example is the leading start of the last recession, reflecting the predictive power of foreign orders. Nevertheless there are also counterexamples, e.g. the development of corporate spread and job vacancies clearly fits the downturn from August 2002 to September 2003, whereas the corresponding specifications do not report a recession. But even if such an intuitive relation is not consistently apparent, it still makes sense to use different leading indicators with respect to different MSARX specifications. The reasons are two-fold: On the one hand, the ex-ante worse specifications can still improve the consensus forecast whenever they contain some independent information from the ex-ante better ones, see Bates and Granger (1969).

[8]Sample 1994:07 - 2010:09 for Tables 2.2, 2.3, 2.4 and 1994:07 - 2008:06 for Table 2.5.

[9]Sample 1999:03 - 2010:09 for Tables 2.2, 2.3, 2.4 and 1999:03 - 2008:06 for Table 2.5. Restricted sample size because of data availability. For instance, the credit growth series is only available from the start of the European System of Central Banks (ESCB) in 1998.

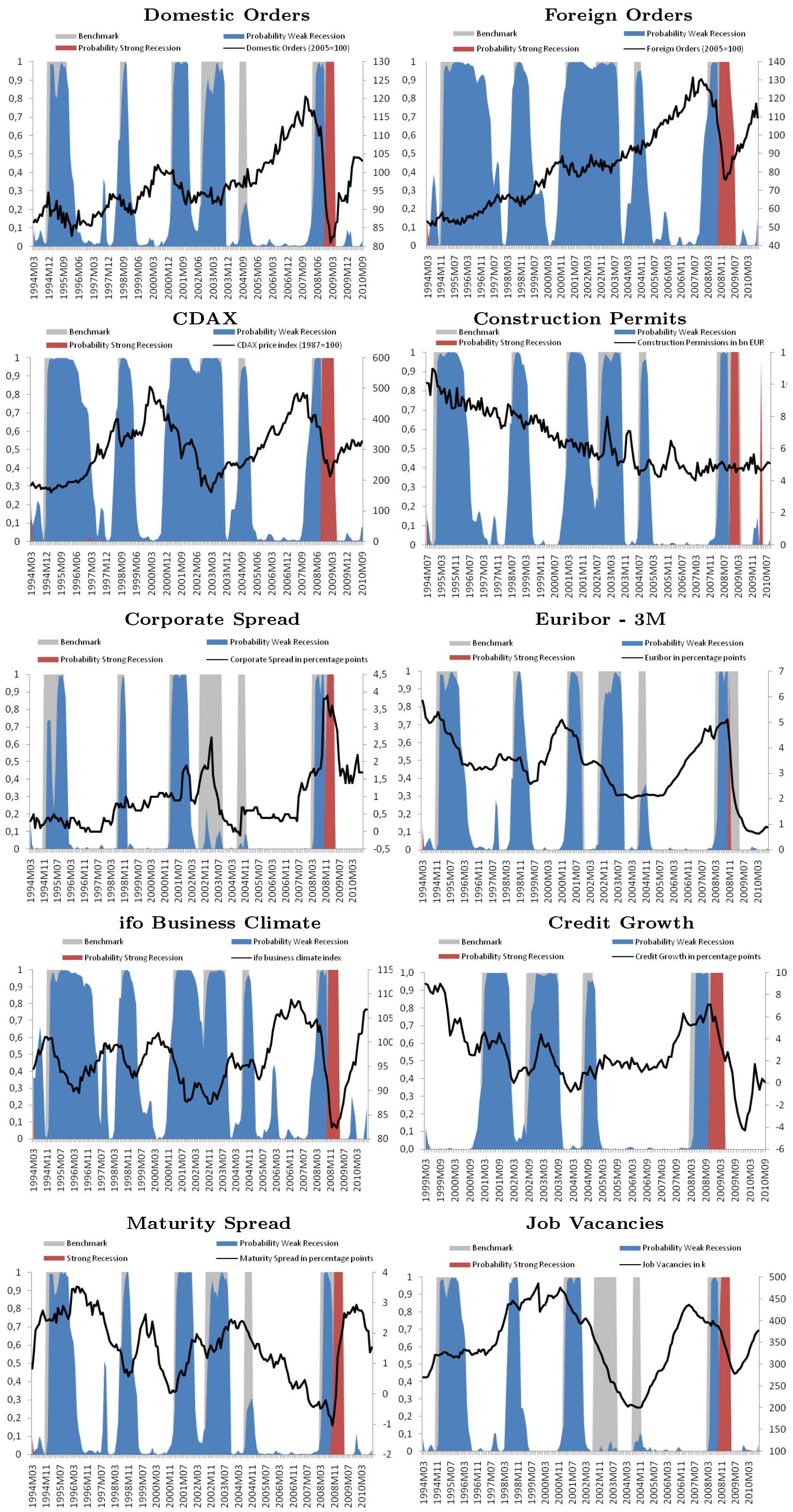


Figure 2.2: In-sample recession state probabilities in bivariate MSARX regressions with 4 regimes (r1 - r4) for the publication 2010:11. The left axis displays the recession probability in each of the regressions, while the right axis shows the course of the embedded exogenous variable, *i.e.* the leading indicator used in the regression. All data is calendar and seasonally adjusted. Expansionary intensities of the business cycle are not distinguished in this figure, but in most cases the strong expansion regime corresponds to the recovery after the financial crisis.

Sample: 1994:03 – 2010:09	recession		expansion		total		$\overline{R}^2$	RMSE	SIC
	#	%	#	%	#	%			
Purely Autoregressive	52	85.25	127	92.03	179	89.95	0.8234	0.0033	-1467.39
Foreign Orders	59	96.72	89	44.72	148	74.37	0.7664	0.0037	-1440.76
Domestic Orders	50	81.97	133	96.38	183	91.96	0.7435	0.0039	-1437.91
Construction Permits [8]	52	85.25	115	85.82	167	85.64	0.8131	0.0033	-1369.53
CDAX	58	95.08	98	71.01	156	78.39	0.7825	0.0035	-1440.37
Corporate Spread	33	54.10	131	94.93	164	82.41	0.7803	0.0036	-1438.46
Euribor - 3M	42	68.85	127	92.03	169	84.92	0.7528	0.0038	-1446.56
ifo Business Climate	58	95.08	91	65.94	149	74.87	0.0980	0.0072	-1447.86
Credit Growth [9]	37	86.05	90	93.75	127	91.37	0.8023	0.0037	- 957.64
Maturity Spread	49	80.33	124	89.86	173	86.93	0.8409	0.0031	-1432.48
Job Vacancies	37	60.66	125	90.58	162	81.41	0.7944	0.0034	-1452.31
AR(1)	-	-	-	-	-	-	0.5916	0.0055	-

Table 2.4: Goodness of fit for bivariate MSARX regressions. The first six columns compare the iterated regime probabilities to the binary benchmark series ($1 =$ recession, $0 =$ expansion) as it is generated by the modified Bry-Boschan dating algorithm. In contrast, the adjusted R^2 as well as the root mean squared error (RMSE) are calculated with respect to the fitted production growth rates. These values outperform results from a simple AR(1) estimation. Finally, the Schwarzian information criteria (SIC) is linked to the likelihood function for each of the regressions.

On the other hand, as was mentioned before, for a new publication a single specification might be subject to a misspecification bias of unknown form, see Timmermann (2006). To a certain extent, this potential weakness can be balanced out in a combined forecast. Finally later on, the OECD Composite Index of Leading Indicators (CLI) will be used as the explained variable in the MS regression which represents another way of information aggregation.
Table 2.4 summarizes the number of correct recession and expansion detections for each of the MSARX specifications in comparison to the Bry-Boschan-type ex-post-dating algorithm mentioned in Section 2.2. In addition to measures related to the regime probabilities, Table 2.2 also reports standard measures linked to the fitted growth rates of the reference series. Those are the adjusted R^2 and the root mean squared error. After regime probabilities are already filtered and the state dependent coefficients are estimated, fitted values can be computed from the different state equations. Therefore each state equation is weighted by the corresponding regime probability. Except for the specification with the ifo Business Climate Index, each of the MS regressions reaches values above 0.7 for the adjusted R^2. In this way they outperform a simple $AR(1)$ benchmark.

In fact the low value in the regression with the Ifo Business Climate might be based on a local maximum found by the expectation-maximization algorithm (see Section 2.3) for the specific publication on November 2010. Thus, this might reveal one of the ten specifications being subject to a misspecification bias, but initially only for this publication. As already mentioned, one way to handle this problem is averaging single forecasts. A similar impression arises after computing the root mean squared error. Again, it is the regression with the ifo Business Climate Index which reaches a far higher value than the others. Table 2.2 also displays the SIC information criterion linked to the log likelihood function of the MSARX estimations. Here none of the results are conspicuously out of range.

2.5 Out-of-Sample Evaluation

2.5.1 Real-Time Forecast Methodology

In general, forecasts with MS models are produced as follows

$$\hat{\xi}_{T+h|T} = \hat{P}^h \hat{\xi}_{T|T}, \tag{2.8}$$

whereby h stands for the forecast horizon. Hence, the future state probability of being in a certain regime comes from the regime probabilities of the last observation. These are weighted by the probability of changing to the certain regime. Obviously, the idea of extending the model to additional regimes, explicitly to a third and to a fourth one, arose during the economic downturn of the financial crisis and the recovery thereafter. But real-time forecasts cannot be based on information received later on. That is why a criterion has been developed which determines when to introduce new regimes on the real-time path. As illustrated in Figure 2.2, the in-sample results show that the probability of the additional regimes has only been allocated since the financial crisis. With real-time forecasts it is clear enough that whenever new regimes are introduced some probability will be allocated to them. The essential question is how high the allocated probability is. The answer to this question lies at the heart of the above-mentioned criterion. With an increasing number of regimes, more extreme events can be reproduced. This enables the model to distinguish between a strong and a weak intensity of the same kind of business cycle phase. Considering a first month as representing a potential outlier, it makes sense to change the number of regimes whenever the probability of the strong intensity exceeds the one of the weak intensity for two consecutive months, *i.e.* with

$t = 1, 2 \ldots$

$$\begin{aligned} \rho\left(S\left(T+t\right) = \text{strong}|y_{(T+t-1)}\right) > \rho\left(S\left(T+t\right) = \text{weak}|y_{(T+t-1)}\right), \\ \rho\left(S\left(T+t+1\right) = \text{strong}|y_{(T+t)}\right) > \rho\left(S\left(T+t+1\right) = \text{weak}|y_{(T+t)}\right), \end{aligned} \tag{2.9}$$

where T+1 stands for the beginning of the out-of-sample forecasts. From an operational point of view, in order to apply this criterion, it is necessary to run both in parallel - the setting with less regimes and the one with more regimes. At this point the question arises whether to increase the number of regimes by one or by two. Here we take a symmetric approach, which in times of a financial crisis means that the total number of regimes can only change from two to four. The reason is quite simple. When introducing a third regime in real-time it will not be clear without laborious computation whether to allocate its probability to a recession or an expansion. Certainly, this advantage of a symmetric approach has to be weighed against the potential instability arising from the additional parameter. At least in our test runs the four-regime model did not provide any evidence of over-parameterization. In the case of two more regimes, two of the four will lead to higher growth rates in absolute values. Hence the regime with the most positive intercept will be allocated to strong expansion, the one with the most negative intercept to strong recession. Then the remaining regimes form the weak expansion and weak recession states.[3] Naturally, this approximation is only feasible whenever the switching intercept is identified as the most relevant part for the iteration of the state probabilities, see Section 2.4.2. Moreover, Welch test results given in Table 2.5 confirm the symmetric approach. Changing the number of regimes due to the above-mentioned criterion should be testable. Hence, we employ a Welch test whose null hypothesis is stated as follows:

> Regimes allocated to different intensities (strong or weak) of the same kind of business cycle phase (recession or expansion) share the same (normal) distribution,

i.e. formally

$$\begin{aligned} H_0 : \left(\mu^{weak}_{expansion}, \sigma^{weak}_{expansion}\right) &= \left(\mu^{strong}_{expansion}, \sigma^{strong}_{expansion}\right) \\ \left(\mu^{weak}_{recession}, \sigma^{weak}_{recession}\right) &= \left(\mu^{strong}_{recession}, \sigma^{strong}_{recession}\right) \end{aligned}$$

[3] In doing so a possible misallocation cannot be excluded categorically since with a single specification there might occur one expansion and three recession regimes or vice versa, but a symmetric approach fits the main empirical findings of the previous sections.

Sample: 1994:03 – 2008:06 Publication: 2008:08	Regime	Switching Intercept	Switching Exogenous Lag	Error's σ	Welch Test Result
Autoregressive	r1	0.0060	-	0.0026	T=8.44
	r2	0.0032	-	0.0010	rejected
	r3	-0.0001	-	0.0012	T=24.73
	r4	-0.0054	-	0.0019	rejected
Foreign Orders	r1	0.0065	0.0041	0.0023	T=18.57
	r2	0.0023	0.0327	0.0018	rejected
	r3	-0.0009	0.0801	0.0002	T=1.97
	r4	-0.0029	0.0242	0.0028	rejected
Domestic Orders	r1	0.0099	0.0423	0.0013	T=21.76
	r2	0.0031	0.0113	0.0027	rejected
	r3	0.0001	0.1650	0.0015	T=21.13
	r4	-0.0041	-0.1187	0.0024	rejected
Construction Permits [8]	r1	0.0113	-0.0889	0.0000	T=17.92
	r2	0.0052	-0.0050	0.0025	rejected
	r3	-0.0015	-0.0317	0.0011	T=2.20
	r4	-0.0027	0.0198	0.0033	not rejected
CDAX	r1	0.0064	-0.0118	0.0025	T=9.63
	r2	0.0039	-0.0099	0.0010	rejected
	r3	0.0002	-0.0042	0.0013	T=19.19
	r4	-0.0052	-0.0061	0.0021	rejected
Corporate Spread	r1	0.0050	0.0263	0.0028	T=12.73
	r2	0.0014	0.0425	0.0017	rejected
	r3	-0.0019	0.0487	0.0024	T=14.72
	r4	-0.0059	-0.0558	0.0012	rejected
Euribor - 3M	r1	0.0065	0.2096	0.0034	T=7.51
	r2	0.0034	0.0128	0.0027	rejected
	r3	0.0026	0.4194	0.0011	T=14.49
	r4	-0.0026	-0.0808	0.0034	rejected
ifo Business Climate	r1	0.0057	0.0805	0.0025	T=15.68
	r2	0.0042	-0.1028	0.0007	rejected
	r3	0.0006	0.1260	0.0015	T=8.07
	r4	-0.0044	0.0378	0.0025	rejected
Credit Growth [9]	r1	0.0075	-0.0485	0.0006	T=4.37
	r2	0.0053	0.0109	0.0021	rejected
	r3	0.0002	0.0029	0.0021	T=29.81
	r4	-0.0073	0.0193	0.0010	rejected
Maturity Spread	r1	0.0041	0.0062	0.0032	T=3.81
	r2	-0.0001	-0.1861	0.0002	rejected
	r3	-0.0018	0.0017	0.0035	T=41.57
	r4	-0.0128	0.1674	0.0001	rejected
Job Vacancies	r1	0.0052	0.0232	0.0031	T=4.56
	r2	0.0034	0.0325	0.0012	rejected
	r3	-0.0006	0.0765	0.0005	T=12.24
	r4	-0.0036	0.0847	0.0022	rejected

Table 2.5: Welch test results for the publication on August 2008. At this point in time the criterion (2.9) suggests a model change from 2 to 4 regimes (r1 - r4). Welch tests are conducted to analyse if this change can be statistically confirmed. The Welch test is related to significant differences in the parameters μ, σ of the normal distributions from different intensities of the same kind of business cycle phase. The hypothesis of identical parameters for different business cycle intensities (weak, strong) can be rejected in almost all cases.

Table 2.5 displays Welch test[4] results for the publication on August 2008, i.e. exactly the point in time after the criterion suggests the real-time change in the number of regimes.[5] As it can be seen, the null can be rejected throughout almost all specifications. Hence, Welch test results support a change to four regimes.

2.5.2 Real-Time Forecasts with the Industrial Production

This section deals with one-month-ahead real-time forecasts with industrial production as the dependent variable. The methodology for how to produce these forecasts including the real-time change in the number of regimes was described in the previous section. Table 2.6 contains measures of forecast accuracy for the single MSARX forecasts and for the combined one (average). Selected measures are the mean absolute value (MAE), the root mean squared error (RMSE) and the Theil coefficient which are widely acknowledged in forecast literature. By definition, Theil coefficients are normalized to the unit interval with 0 representing the perfect fit. In this context, Timmermann (2006, p.1) mentions that 'simple combinations that ignore correlations between forecast errors often dominate more refined' ones.

Yet, the average, as listed in Table 2.6, does not only achieve the best values because of the averaging effect, but also because of the fact that the model changes the number of regimes due to the criterion from Section 2.5.1. This change takes place with the forecast for September 2008.[6] One reason for the Theil coefficient not reaching lower values[7] is the slightly delayed recession detection, also see Figure 2.3.

[4]Taking S_t as a filtration calculus of the conditional expectation shows how to approximate parameter μ by the intercept and the exogenous parts of the equation:

$$\begin{aligned}
E\left(y_t|S_t\right) &= E\left(E\left(y_t|S_t\right)|S_{t-1}\right)\\
&= E\left(\beta_0^{S_t} + \beta_{1,y}^{S_t} E\left(y_{t-1}|S_t\right) + \beta_{1,x}^{S_t} x_{t-1}|S_{t-1}\right)\\
&= \beta_0^{S_t} + \beta_{1,x}^{S_t} x_{t-1} + \beta_{1,y}^{S_t} E\left(y_{t-1}|S_{t-1}\right)\\
&= \beta_0^{S_t} + \beta_{1,x}^{S_t} x_{t-1} + \beta_{1,y}^{S_t}\left(\beta_0^{S_{t-1}} + \beta_{1,x}^{S_{t-1}} x_{t-1} + \beta_{1,y}^{S_{t-1}} E\left(y_{t-2}|S_{t-2}\right)\right).
\end{aligned}$$

Since for each β-term one has $\beta << 1$, the last term consisting of products of βs can be neglected. A similar assessment identifies the error's σ^2 as the essential part of the conditional variance.

[5]Note that with respect to criterion (2.9) such a change in the number of regimes is not suggested for any other publication (vintage) before.

[6]The title of the Gemeinschaftsdiagnose, the professional opinion of important German economic research institutes, in fall 2008, was 'Germany on the edge of a recession'. In contrast to this report, the real-time introduction of new regimes for September 2008 could be interpreted as Germany no longer being on the edge, but in the middle of a recession becoming deeper than all previous recessions.

[7]Another reason is that, compared to the binary series, the averaged MS state probabilities are likely to cover only a certain range between 0 and 1. Thus, the measures of forecast accuracy could only be improved by generating a binary variable out of the MS state probabilities according to the 0.5 threshold.

Sample: 2007:02 – 2010:12	MAE	RMSE	Theil
Average	0.2948	0.3898	0.4060
Autoregressive	0.3070	0.4439	0.4309
Foreign Orders	0.3651	0.5007	0.4340
Domestic Orders	0.3141	0.4385	0.4203
Construction Permits	0.3050	0.4394	0.4112
CDAX	0.3234	0.4569	0.4421
Corporate Spread	0.3609	0.4958	0.4921
Euribor - 3M	0.3132	0.4485	0.4537
ifo Business Climate	0.3564	0.4991	0.4860
Credit Growth	0.2559	0.4395	0.4128
Maturity Spread	0.2987	0.4585	0.4279
Job Vacancies	0.3774	0.5038	0.4564

Table 2.6: Measures of forecast accuracy for the real-time forecasts from bivariate MSARX specifications and from the composite approach. Averaging the forecasts and installing a real-time change in the number of embedded regimes due to criterion (2.9) improves the out-of-sample performance.

Compared to the recession start in March 2008, as it is reported by the benchmark model in October 2008 (7 months later), the MS recession probability exceeds the 0.5 threshold in August 2008. On the one hand, this represents a delay of 5 months with respect to the observational time. On the other hand, the recession is recognized significantly earlier by the MS model than by the benchmark method (publication time). As the forecast for August is made in July (one month forecasting horizon), the time in advance between the MS and the benchmark model stands at 3 months. Additionally, the recession probability forecast for July 2008 is above 30%. Such a careful indication of a recessionary period cannot be provided by a binary detection decision (1 = recession, 0 = expansion). Both MS and the Bry-Boschan-type ex-post-dating algorithm continuously announce the last recession between August 2008 and April 2009. Compared to the end of the recession in April 2009, as it is reported by the benchmark series in November 2009, the MS forecast for July 2009 is the last above the 0.5 threshold (delay of 3 months). Again considering the point in time when the information about the end of the recession was provided, the MS model is 4 months earlier. Thus the forerun of the MS is even longer in the case of the end of the recession than the beginning.

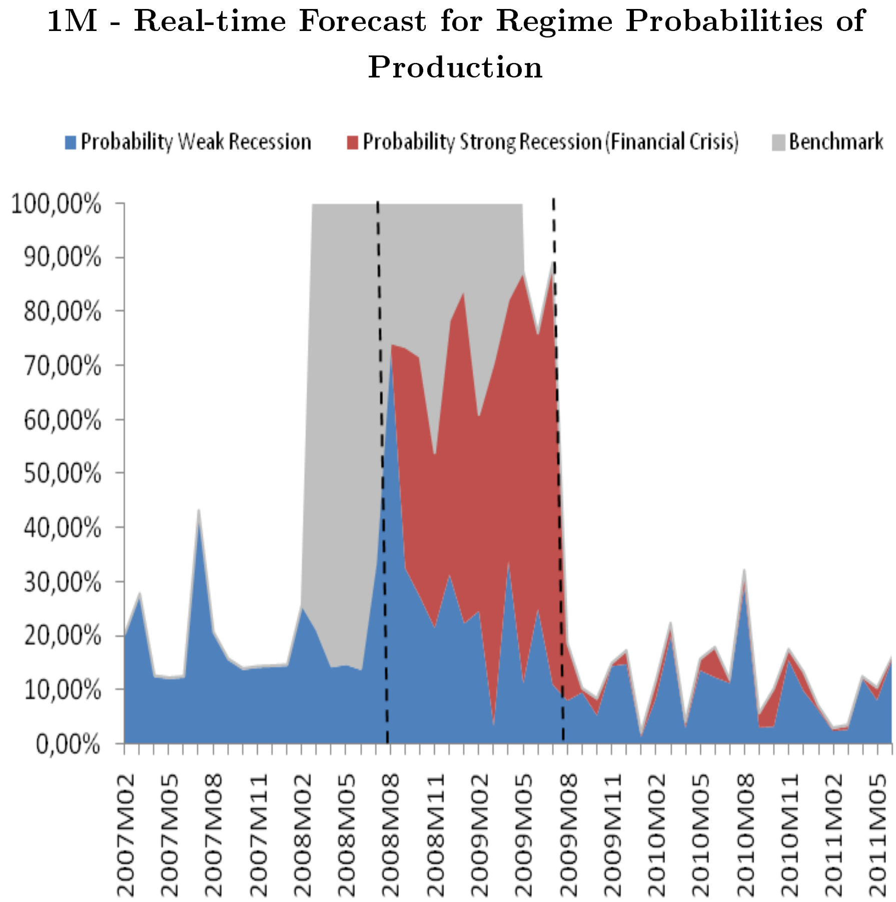


Figure 2.3: The regime probabilities are averaged one month ahead real-time forecasts of the different MSARX specifications. Probabilities of weak and strong recession intensity are added up to a total recession probability. For September 2008, the model changes from 2 to 4 regimes due to criterion (2.9), where the new regime clearly points to the magnitude of the economic decline. When defining a recession on the 0.5 threshold, the downturn is predicted continuously between August 2008 and July 2009 (dashed lines). Actually this represents a delay to the recession start, as it is reported later on by the ex-post dating algorithm. But considering the point in time when the recession is recognized, the MS model is 3 months earlier than the benchmark method.

Together with the aggregated recession probability not reaching values above 90 percent, this reveals a certain restraint towards an erroneous recession declaration, which seems to be functional with respect to the forecast accuracy. Indeed there is no extra period in the out-of-sample evaluation where the economic situation is misinterpreted as a recession.

2.5.3 Real-Time Forecasts with the OECD Composite Leading Indicator

Although the MS model outperforms the ex-post-dating algorithm, the question arises as to whether there was any alternative in recognizing the recession in advance. Lahiri and Wang (1994) employ another MS model to the U.S. Composite Index of Leading Economic Indicators (CLI). Such monthly data for Germany is provided by OECD (2011).

The idea behind the CLI is to generate a synthetic series that represents a lead to the business cycle and anticipates its turning points. To achieve this, leading indicators - similar to the way they are used as regressors in the previous estimations - are aggregated. That is why the CLI is also subject to revisions. Before aggregation, the data is seasonally adjusted, outliers are eliminated, trends are removed, and filters for smoothing and normalization are applied in order to obtain homogenized cyclical amplitudes for each of the component series. It is not the topic of this chapter to discuss the OECD methods in detail, but it turns out that the procedure above leads to the CLI often exhibiting relative undecidedness between up- and downturns at the current edge. Nevertheless with the Hamilton filter generating the state probabilities endogenously out of the observations, and with the result of lagged recession recognition in the case of the industrial production, it is quite appealing to run a specification where the above-mentioned reference series is substituted by the OECD CLI. In doing so, some differences to the previous MS regressions have to be considered. Firstly, smoothing backwards by a moving average is no longer necessary since the series is already smoothed. Secondly, the lag choice, described in Section 2.4.2, only makes sense for a purely autoregressive specification since there must be a bias with leading indicators standing on both sides of the equation in a different manner (aggregated versus disaggregated). In fact the lag choice results in favoring no autoregressive terms, so that the right hand side of the equation only consists of a switching intercept and an error term. As a consequence of this parsimonious design it is not sufficient to choose the regimes with the lowest intercepts to stand for recessions, but to request these intercepts to be negative. In Figure 2.4 the MS regression with the OECD CLI delivers an early signal for the recession linked to the financial crisis, but forecasts are very volatile. Among the real-time out-of-sample predictions between February 2006 and June 2011 there are three periods (8 months), in which the economic situation is misinterpreted as a recession.

1M - Real-time Forecast for Regime Probabilities of CLI

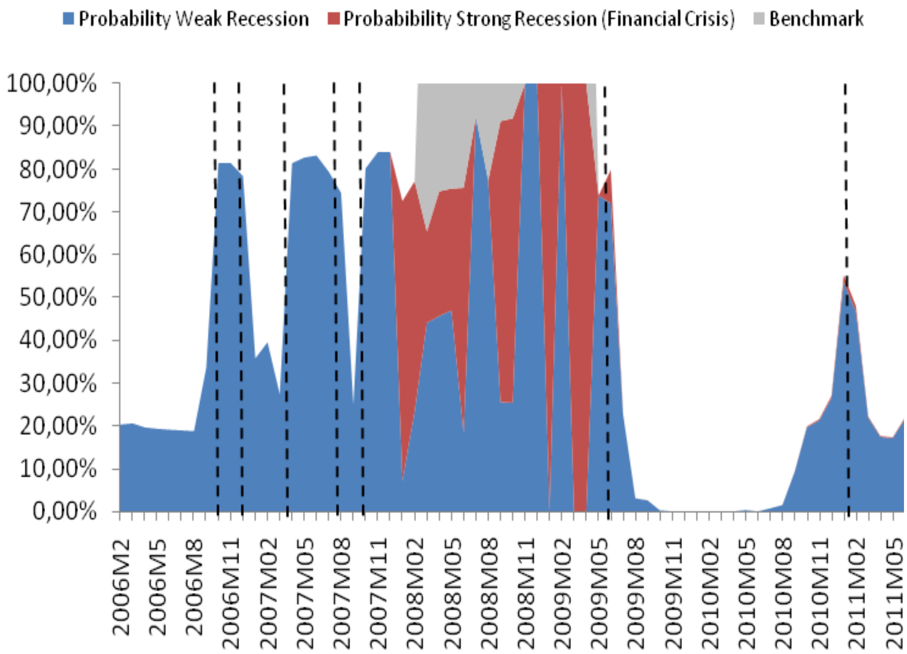


Figure 2.4: The regime probabilities are one-month-ahead real-time forecasts with no autoregressive terms. Probabilities of weak and strong recession intensity can be added up to a total recession probability. For December 2007 the model changes from 2 to 4 regimes according to the criteria described in section 2.5.1, where the new regime clearly points to the magnitude of the economic decline. When defining a recession on the 0.5 threshold, there are three periods (October 2006 - December 2006, April 2007 - August 2007 and January 2011), in which the economic situation is misinterpreted as a recession. Among the very volatile forecast results an early indication of the approaching recession can be identified in October 2007. Compared to the beginning of the recession in March 2008 as is reported by the benchmark series, this early signal would represent a lead of 6 months. The recession is then predicted continuously between October 2007 and June 2009 (dashed lines).

The early signal is given six months in advance, whereas a timely signal in accordance with the forecast horizon would have to come one month ahead of the publication point. This reveals a general problem with CLI data: Given the high number of misinterpreted recessionary phases it is not clear whether the forerun is stable and always equal to 6 months as is claimed by the OECD. Nevertheless, from an operational point of view, MS regressions on the CLI and on the industrial production can be complementary. While the former might signal the recession in advance, the latter can confirm it accurately a shorter time after its onset as is usually needed by official chronology decision processes like the NBER turning point dating procedure.

2.6 Conclusion

This chapter uses a Markov Switching framework applied to German monthly real-time data. While the appropriateness of the method for business cycle applications is well-known since Hamilton's innovation in 1989, based on current literature there are some new insights which can be fruitfully applied to monthly German real-time data: Given limited data records, it is appealing to connect Timmermann (2006)'s notion of a single forecast being subject to a misspecification bias with the Markov Switching model generating each of the single forecasts. In order to reduce the bias, forecasting results are averaged. When generating the forecasts as above, several macroeconomic and financial leading indicators serve as exogenous variables in bivariate MSARX regressions.
Allowing the MS model to change the number of embedded regimes in real-time stabilizes forecasting results. By introducing a criterion for the real-time regime change it is also possible to determine the point in time from which the recession after the financial crisis structurally exceeded the previous ones. In our analysis this turns out to be September 2008, where the forecast is made in August - one month before the investment bank Lehman Brothers declared bankruptcy. The MS real-time forecasts from February 2007 to Mai 2011 for industrial production outperform a modified non-parametric ex-post-dating method based on the work of Bry and Boschan (1971) as well as Harding and Pagan (2002), while revealing similar characteristics: On the one hand recession start and end are recognized too late, while the delay for the end of the recession is considerably shorter. On the other hand, at least when considering different specifications and changing the number of regimes, no business cycle phase is misinterpreted as a recession. In order to counterbalance the aforementioned inertia of the MS model in the case of industrial production, it is appealing to apply it to the OECD Composite Index of Leading Indicators, which seeks to be a leading proxy for the business cycle. In doing so the findings for the combination of the MS model and a leading index are more ambivalent. On the one hand, we can confirm Lahiri and Wang (1994)'s result of obtaining an early signal. On the other hand, predictions are of low quality since in several instances recessions are mistakenly declared. Finally several extensions of our MS framework are possible, such as the inclusion of different forecast horizons.

Chapter 3

Agent-based Risk Management - A Regulatory Approach to Financial Markets

3.1 Introduction

In the aftermath of the financial crisis there has been a broad societal consensus on the need to improve financial regulation. Although different actions have been undertaken worldwide - the Basel III (2010) regulations, the Dodd-Frank Act and the creation of the European Systemic Risk Board (ESRB) being the most significant - doubts remain as to whether these actions have been sufficient (Chappe and Semmler, 2012, for a survey). One of the relatively unaffected fields are the risk management models, central to the first pillar of Basel II, which allow, alongside a standard approach, the application of the so-called internal model method (IMM). The idea behind this value-at-risk (VaR) approach is to install an effective risk calculation inside banks, influencing daily business strategy and creating an anchor for supervision by national controlling authorities. In the end, the value-at-risk determines for many banks the essential part of the regulatory capital, which is supposed to cover different risks.
On the one hand, the new regulatory framework by the Basel Committee on Banking Supervision (2010), Basel III, tries to draw lessons from the financial crisis by changing the conditions for what is accepted as regulatory capital. In particular higher core capital quotas are required and countercyclical buffers

are created. On the other hand, little is done to improve the concrete risk management models linked to the first pillar of Basel II. Systemic risk should be accounted for at a point where risks are concretely quantified, namely in the VaR approaches for market, credit and operational risk, instead of shifting it towards the more abstract rules of the second pillar[1]. This is the starting point for the paper at hand and we will show that many of the proposals for improving financial regulation are reflected in the VaR results of agent-based simulations.

Like in other areas of economics there is a huge need for a turnaround in financial market modeling in order to make amends for the mistakes that initiated the recent crisis. Yet one could posit that, at least in academic financial literature, half of the ground is covered. For instance Shiller (2003) already described a development 'From efficient market theory to behavioral finance'. Regrettably there is no parallel within risk management models, where the assumption of the no-arbitrage condition prevails. Although the limitations of (no-)arbitrage are well discussed in the literature, *e.g.* compare Shleifer and Summers (1990) and Ritter (2003), the innovation in risk modeling is taking place - if at all - very slowly. Colander, Föllmer, Haas, Goldberg, Juselius, Kirman, Lux and Sloth (2009) provide an excellent analysis of the 'the financial crisis and the systemic failure of academic economics' which, in their view, was caused by 'a misallocation of research efforts in economics.'[2] The same seems to be true for the special field of risk modeling although alternatives for financial market modeling have been around for some time, see DeLong, Shleifer, Summers and Waldmann (1990), Campbell and Kyle (1993) and Arthur, Holland, LeBaron, Palmer and Tayler (1996) among others. Nevertheless, their analysis, in which market participants partly follow rules of thumb or market psychology, was widely ignored in risk management models although staff members with trading experience should have known better.
Besides the assumption of no-arbitrage, which has its most popular application in the formula of Black and Scholes (1973), there are further hypotheses underpinning these models which have become ever more empirically questionable. One implicit cornerstone of current risk modeling are rational expectations, which *ex ante* imply a centered stochastic error term capturing the difference between the representative agent's [3] expectation and the current equilibrium

[1]See Basel Committee on Banking Supervision (2006) for an exact definition of the pillars.
[2]Colander et al. (2009, page 1) additionally state: 'We trace the deeper roots of this failure to the profession's insistence on constructing models that, by design, disregard the key elements driving outcomes in real-world markets. The economics profession has failed in communicating the limitations, weaknesses, and even dangers of its preferred models to the public.'
[3]The question, as to whether the assumption of a representative agent is feasible, has been

price. But based on the study of financial market time series, the assumption that market participants never make systematic mistakes in terms of their expectations seems to be questionable.
Up to the present, internal models of large banks are mainly driven by the assumption that the logarithmic risk factors of their financial instruments follow more or less a random walk. This kind of modeling is rooted in Fama (1970)'s efficient-market hypothesis, proposing that at least all publicly available information is already captured by current prices. Following this logic, changes in future prices due to changes in the underlying process of information would be well enough assessed by random disturbances. Implicitly, this puts market prices close to a valuation which is mainly based on changes in the fundamentals. The problem with this perspective is, as Lux (1998) points out, that observable financial market phenomena often do not coincide with the process of new information production. For instance, turbulent market periods, producing fat-tails in the distribution of returns, often cannot be explained sufficiently by a change in the fundamentals according to the upcoming news about a company.[4]

Meanwhile, a growing academic financial literature develops models[5] that can reproduce the so-called market anomalies[6], in particular including the phenomena of excess volatility and volatility clustering - both often linked to the origination of bubbles. However, these models are still under-represented in risk management, though, as this paper shows, they are able to work under the regulatory requirements. In particular, agent-based models are rarely used for risk management although they provide some useful insights about the underlying factors producing the anomalies.[7] In the present paper we compare the performance of an agent-based approach with the standard VaR approach

raised by Kirman (1992) among others. The model presented here deals with heterogeneous agents for each market.

[4]Another feasible approach in this context is to make the information process more complex and more realistic (Park and Sabourian, 2011). In general, the literature distinguishes between rational and irrational herding as well as rational and irrational bubbles. As discussed at the end of this paper, the presented regulatory approach might also be implemented with a rational bubble model. Anyway, we start to employ an irrational one.

[5]Such models are GARCH/FIGARCH models (Bollerslev, Engle and Nelson, 1994; Baillie, Bollerslev and Mikkelsen, 1996, among others), FRACTAL models (Mandelbrot and van Ness, 1968; Mandelbrot, Fisher and Calvet, 1997, among others) and Agent-based Computational Economic (ACE) models (Lux, 1998; Lux and Marchesi, 2000, among others), with which this paper deals in detail.

[6]In this context it might be more appropriate to speak about stylized facts, see Pagan (1996). Appendix 3.5 contains some more information on financial market time series properties highlighting German stock exchange data before and during the last financial crisis.

[7]While we focus on the agent-based approach, the next proposal for a 'framework for more resilient banks and banking system' (Basel IV) should, conceivably, encompass different approaches within the regulatory framework so that 'model uncertainty (can be) taken into account by applying more than a single (type of) model', see Colander et al. (2009, page 6).

and a GARCH risk estimation.
This paper contributes to the literature by making the idea of agent-based computation compatible with the existing regulatory framework. For the purpose of risk calculation, a modified version of the Lux and Marchesi (2000)-model is developed. In detail, the new model employs an inhomogeneous Poisson process to establish the link between expected excess demand and price realizations and includes an estimation of model-adjusted trading volumes. Extensions are two-sided: On the one hand, the modifications facilitate a meaningful calibration to bring the agent-based simulations closer to the data. This enables the model to work within the Basel framework, where for practical reasons [8] usually only the last 250 realizations of risk factors are taken into account. On the other hand, the model is extended to include the derivatives market. The main reason for this extension is the agent-based price process of the underlying asset reflecting a recurring alternation between non-turbulent and turbulent market periods, as described by Chen, Lux and Marchesi (2001). If, in addition, there are derivatives, which can generate a feedback effect on the price of the underlying asset, this will increase the frequency of turbulent markets (bubbles). In this way, the extension reveals a key ratio of the derivatives market which in the future, if monitored by national regulatory authorities, may lower the influence of speculative derivatives trading.
The remainder of this paper is structured as follows: Section 3.2 presents the modified agent-based model, in detail for the stock market in Sections 3.2.1 and 3.2.2, while Section 3.2.3 introduces the derivatives market including its feedback mechanism on the underlying asset. For comparison, Section 3.2.4 deals with the data generating processes from existing internal model methods: First, the still prevailing Black-Scholes-Merton type models and in addition risk estimation with a standard GARCH approach. Section 3.3 discusses the methods of calibrating these models. Section 3.4 presents numerical results based on German financial data before and during the last financial crisis. Those results are evaluated with respect to the regulatory objective of future market stability. Section 3.5 concludes. Supplementary material is provided in an Appendix.

3.2 A Regulatory Agent-Based Model

3.2.1 Heterogeneous Agents in a Dynamic System

What lies at the heart of the model by Lux (1998) and Lux and Marchesi (2000) are three different classes of agents: (1) optimistic noise traders (+), who want

[8] The daily risk of bank portfolios needs to be computable during a night batch.

to buy the stock out of a motivation that will be clarified in detail, but is, at least based on a second-round effect, part of the model [9]; (2) pessimistic noise traders (-), who want to sell the stock out of a motivation opposite to the previous one; and (3) fundamental traders (f), who want to make money by arbitrage between the market price and a fundamental price which arises from a CAPM computation. Optimists and pessimists as a whole form the group of noise traders, i.e.

$$n_t^+ + n_t^- = n_t^n. \tag{3.1}$$

The number of market participants is then given by

$$n_t^n + n_t^f = N. \tag{3.2}$$

At each point in trading time the agents interact and decide whether to stay or to change class.

Noise traders should not be thought of as non-informed traders because they use *a priori* the available information to calculate the (expected) difference in returns in order to decide whether to stay with the noise trading or to become a fundamental trader. Whenever there are more noise traders than fundamentalists *a posteriori*, this means that the majority of market participants has classified fundamental valuation as subordinate to other factors like the current price trend. However, such a decision in favor of non-fundamental factors is often initialized randomly by stochastic terms also used in the agent-based model. In that case, the first impulse for (more) noise trading comes from outside the model and is thus not far from the characterization of Park and Sabourian (2011) quoted in footnote 9.

Before turning to the mathematical details, the rationale of the model can be described as follows: At the beginning (and recurrently in between), the number of agents representing optimists, pessimists and fundamentalists is uniformly distributed and the price is close to the fundamental value determined by CAPM. By means of a stochastic impulse from outside the model, the price increases so that a positive price trend arises. *Ceteris paribus* , this leads to a slight excess of optimistic agents creating a corresponding excess demand. In the next period, this excess demand will cause an even more positive price trend, if not balanced by arbitrage transactions. Whenever arbitrage transactions are not strong enough to lead prices back, this creates a self-reinforcing process (herding behavior) that will not be interrupted before a large deviation between market and fundamental price has arisen. An (irrational) bubble has been originated.

[9]In contrast, there are also models in the literature, in which 'noise traders have no information and trade randomly. These traders are not necessarily irrational, but they trade for reasons not included in the model, such as liquidity' (Park and Sabourian, 2011).

For the purpose of risk calculation, according to the Basel framework, we remodel the design by Lux and Marchesi (2000). In this section, we describe how the number of agents in each trading class may change over time. These dynamics are driven by rules of thumb or behavioral heuristics.[10] In particular, we deal with the question of what determines the switch between the types of traders. The transition process describing the agents' dynamics is assumed to follow a first order inhomogeneous Markov chain, which has a uniform starting distribution and whose transition matrix looks like

$$\begin{pmatrix} 1-\pi_{+-}^t-\pi_{+f}^t & \pi_{+-}^t & \pi_{+f}^t \\ \pi_{-+}^t & 1-\pi_{-+}^t-\pi_{-f}^t & \pi_{-f}^t \\ \pi_{f+}^t & \pi_{f-}^t & 1-\pi_{f+}^t-\pi_{f-}^t \end{pmatrix}. \tag{3.3}$$

We will discuss two of the transition probabilities in detail and refer to the Appendix 3.5 for the rest. The first one is the transition among the noise traders - here from the optimist (+) to the pessimist (-):

$$\pi_{+-}^t = \min\left(v_1 \underbrace{\frac{n_t^-}{N}}_{(1)} \exp\left(-\left(a_1 \underbrace{\frac{n_t^+ - n_t^-}{n_t^n}}_{(2)} + a_2 \underbrace{\frac{p'(t)}{p(t)}}_{(3)}\right)\right), 1\right). \tag{3.4}$$

In equation (3.4) three different influencing factors are emphasized: (1) denotes a weighting factor, here equal to the pessimists' fraction of all market participants. Its meaning will be clarified using the example of the second transition probability. (2) denotes the impact of herding behavior. In the case of a lower number of optimists this will reinforce the change to the pessimists' class. (3) denotes the impact of extrapolating the recent price trend. Here a positive price derivative will damp the change to the pessimists' class. $a_1 = 1 - a_2$ controls the weighting between majority opinion and actual trend. We choose a uniform weighting due to a lack of information indicating that another weighting would be more appropriate. $v_1 = 1 - v_2$ sets the frequency of revaluation for the transitions within the class of noise traders. v_2 is used for the transitions from a noise trader to a fundamentalist and vice versa. We choose slightly different values ($v_1 = 0.6, v_2 = 0.4$) because we assume the interaction process in the first case to be faster than in the second. The reason therefore is that noise and fundamental traders in principle believe in different investment strategies [11] and it takes some time to get to know the other side before one can opt for a change.

[10] One insight of behavioral economics is that people tend to overestimate the influence of the most recent price trend (Kahnemann and Tversky, 1973). Other insights like non-transitivity clearly contradict the standard set of microeconomic preferences.

[11] Fundamental traders basically believe in the efficient-market hypothesis. Noise traders for instance can be thought of as being chartists.

The second transition probability described here is responsible for the change from the pessimistic agents' class to the fundamental one:

$$\pi_{-f}^{t} = \min\left(\underbrace{v_2 \frac{n_t^f}{N}}_{(1)} \exp\left(-a_3^{t,e}\left(\underbrace{R - \frac{d_t + \frac{1}{v_2}p'(t)}{p(t)}}_{(2)} - \underbrace{\left|\frac{p_t^f - p(t)}{p(t)\left(1+\hat{r}_M\right)}\right|}_{(3)}\right)\right), 1\right). \tag{3.5}$$

Again the dynamics are restricted to three factors: (1) denotes the corresponding weighting factor from above. To illustrate its purpose, define the driving force $\tilde{\pi}_{-f}^{t}$ so that $\pi_{-f}^{t} = \frac{n_t^f}{N}\tilde{\pi}_{-f}^{t}$. In that case the number of fundamentalists is

$$n_{t+1}^f = \frac{n_t^f}{N}\left(\tilde{\pi}_{-f}^{t} n_t^- + \tilde{\pi}_{+f}^{t} n_t^+\right) + \pi_{ff}^{t} n_t^f \tag{3.6}$$

The probability of becoming a fundamentalist thus corresponds only to a proportion of the driving forces. The intuition behind the weighting factor is the variety of a decelerator effect. Suppose the impulse for changing to the fundamentalists (the driving force) is high in a situation in which at the same time only a few fundamentalists have remained in the market (close to the peak of a bubble). Then the driving force can act only in part (high decelerator effect) which explains why a bubble can continue to grow or does not burst too abruptly. On the other hand, suppose the impulse for changing to the fundamentalists (driving force) is relatively low, while a substantial part of market participants follows the fundamentals. Then the existing driving force can act almost fully (low decelerator effect) which prevents that bubbles continuously arise. Both effects are consistent with real-world financial market behavior. Moreover the weighting factor helps to guarantee that the transition matrix from above really describes a probability measure, *i.e.* row-wise for instance $\pi_{+-}^{t}, \pi_{+f}^{t} < 1$.

(2) denotes the return of a pessimist, which itself consists of the return of an alternative investment (R) less the return of the stock. d stands for the dividend. The pessimist extrapolates the recent price trend, which - if negative - encourages him to sell. (3) denotes the return of a fundamentalist and corresponds to the discounted profit of arbitrage between the fundamental price p^f and the market price p. Altogether, a change is motivated by the comparison of the returns, where additionally $a_3^{t,e}$ measures the traders' reaction to the return differences. When looking for dynamics in this reaction, adaptive expectations tell us

$$a_3^{t,e} = a_3^{t-1,e} + \lambda\left(a_3^{t-1} - a_3^{t-1,e}\right),\ a_3^{1,e} \doteq 1, \tag{3.7}$$

where we treat the second summand as a forecast error with $\epsilon_t^{a_3} \sim N\left(0, \left(U\left[0, 0.01\right]\right)^2\right)$. The small standard deviation [12] ensures that the reac-

[12] U denotes here the drawing of uniformly distributed random numbers within the interval.

tion on return differences is kept within a realistic range, *e.g.* $]0,2]$ can either mean that the difference in returns almost does not matter, or that its importance is doubled. Modeling $a_3^{t,e}$ in the way mentioned above opens space for interpretation instead of arbitrarily setting the parameter unequal to 1, as is partly done in Lux and Marchesi (2000). For instance let us consider the change from a noise trader to a fundamentalist: Firstly if $a_3^{t,e} > 1$, then the change is reinforced. This might reflect adaptive learning. In that case, noise traders start studying financial literature, which depending on the literature can reinforce a change to the fundamentalists. Secondly if $a_3^{t,e} < 1$, the change is weakened. Noise traders might be overconfident with respect to their investment strategy and remain unchanged (Alicke and Govorun, 2005), although in the meantime a significant deviation between the fundamental and the market price has arisen which, if not $a_3^{t,e} < 1$ would create a stronger impulse for changing to the fundamentalists' class.

In a nutshell, the objective of all market participants is to maximize (short-term) returns. To achieve this goal, traders sometimes follow relatively sophisticated models like the Capital Asset Pricing Model (CAPM) to find the fundamental value of an asset and to benefit from arbitrage (fundamental trader). Another time, they rely on simple behavioral heuristics (noise trader). The model entails non-stable preferences and for simplicity reasons abstracts from wealth dynamics.[13] Preference consistency would be a desirable feature, but there is often a trade-off with model performance to describe observed market anomalies. Non-stable preferences are a viable tool to cope with these stylized facts. It is not the purpose of risk management models to form expectations about future prices under preference-consistency, but to simulate what can be the p% worst tomorrow's price return.

So far, we covered the agents' dynamics. The missing parts are the fundamental price and the question of how the agents' dynamics translate into a market price. But before dealing with these issues in detail it is worth highlighting the trading time underlying the price process. The model uses discrete time linked to intraday trading. Every twelve minutes [14] trading takes place, the agents decide to which class they want to belong and stock prices are updated.

[13]Agents make no profit taking and there is no market exit by bankruptcy. Against the background of an in-depth calibration of the model, including wealth dynamics goes beyond the scope of this paper.

[14]We count 8 hours for a trading day. Hence an elementary time unit is equal to 0.025 trading days.

3.2.2 Price Process

Fundamental Price Process

In principle the fundamental price originates from a discounted cashflow model and, in analogy to simple neoclassical economic models, the fluctuations just emerge from shocks to the (infinite) dividend cashflow and the corporate growth rate. Fundamental traders use this price to identify arbitrage opportunities. In contrast to Lux and Marchesi (2000), in this paper we utilize the Capital Asset Price Model (CAPM) and Gordon's formula to compute the fundamental price. From the perspective of a modeler this does not mean concealing all empirical doubts which arose over the years with these models, but just assuming that there are market participants who really believe in them [15]. More precisely, they believe that market prices sooner or later converge to the CAPM results. Otherwise it would not make sense to employ them for the search of arbitrage opportunities. The main reason why to use the CAPM for the determination of the fundamental price is that it allows for an automatized procedure and that it is still widely acknowledged despite the above-mentioned empirical concerns.[16] In addition, when looking back to the transition probability in equation (3.5), the requirement to determine the return of the alternative investment R, the discount rate $\hat{r}_M$ and the fundamental price p^f becomes obvious. While selecting a stock index as the 'market portfolio' (M) [17], the CAPM computes individual β-factors and risk adjusted rates by

$$\hat{\beta}_i = \frac{cov\left(\hat{r}_M, \hat{r}_i\right)}{\sigma_M^2}, \quad \hat{r}_f^i = \frac{\hat{r}_M \hat{\beta}_i - \hat{r}_i}{\hat{\beta}_i - 1}, \quad i = 1, \dots, 30. \tag{3.8}$$

Then the average of those rates is selected to stand for the return of the alternative, i.e.

$$R := \frac{1}{n} \sum_i^n \hat{r}_f^i, \quad n = 30. \tag{3.9}$$

Simultaneously the return on the market portfolio is chosen as the relevant discount rate for arbitrage activity. [18] Moreover $(1-b)$ stands for the observable payout ratio per share and $d_t = (1-b)\, g_t$ for the dividend per share. We add uncorrelated shocks to the profits per share (g) and to the core growth rate

[15] In view of finance lessons at universities this would not be too unrealistic.

[16] Of course, many other ways are conceivable to gain information about $\left(\hat{r}_M, R, p_f\right)$, for instance by a survey of financial analysts. Unfortunately such data is expensive to obtain and results are very subjective. However, there is no reason to expect substantially different risk results from other fundamental price information.

[17] Data for all stocks listed in the German stock index DAX is loaded for the agent-based risk calculation.

[18] Because of the well-known risk that the trader fails to make profit by arbitrage, the $\hat{r}_M$ should be higher than R.

(cg) of the company value [19] to allow at least for a minimum variation of the fundamental price, i.e.

$$\begin{aligned} g_t &= g_1 + \epsilon_t^g, \qquad \epsilon_t^g \sim N\left(0, \sigma^g\right) \\ cg_t &= cg_1 + \epsilon_t^{cg}, \qquad \epsilon_t^{cg} \sim N\left(0, \sigma^{cg}\right). \end{aligned} \tag{3.10}$$

Finally Gordon's formula, basically the infinite geometric series, delivers the fundamental price by

$$p_t^f = \frac{g_t\left(1-b\right)}{R - b\ cg_t}. \tag{3.11}$$

In this context, Campbell and Kyle (1993) call their fundamental traders 'smart money', since those traders use a more sophisticated procedure for their price expectation compared to the noise traders' behavioral heuristics. But, in general, this does not guarantee them better returns. Despite their supposed superiority there is much evidence that in periods with strong price gains or losses they give up their position to follow the trend (herding behavior) and become noise traders, see Shleifer and Summers (1990).

Market Price Process

So far, the number of agents in each trading class and for each point in time has been calculated. Now the question arises as to what conclusions for the market price can be drawn from this. Lux and Marchesi (2000) define the (expected) excess demand ((E)ED) of noise traders by

$$EED_t^n = \left(n_t^+ - n_t^-\right) vol_{model}. \tag{3.12}$$

Thus, in the case of a majority of optimists versus pessimists, the EED will be positive, when not the case the EED will be negative. Simultaneously, the excess demand corresponds to the difference between the noise trader groups weighted by an average trading volume per transaction vol_{model}. We will discuss in Section 3.3 why to use the subscript *model* in there. The expected excess demand generated by fundamentalists is then defined by

$$EED_t^f = \frac{p_t^f - p\left(t\right)}{p\left(t\right)}\ n_t^f\ vol_{model}. \tag{3.13}$$

Here a higher fundamental than observable price is an incentive for the 'smart money' trader to buy, or to sell given a lower fundamental price. This incentive is divided by the current market price to form a fraction of the fundamentalists' class as a whole. [20] This subset, weighted again by the average trading volume per transaction, corresponds to their expected excess demand. The total

[19] σ^g and σ^{cg} are intuitively set to 1 cent respectively to 100 basis points as there is no data available on which to base more refined values.

[20] A possible explanation is that the rest of the fundamental traders could clear their transactions or are not willing to trade at all.

expected excess demand for a certain point in time is then given by

$$EED_t = EED_t^n + EED_t^f. \tag{3.14}$$

It will be readily apparent why, in contrast to Lux (1998), we explicitly refer to an expectation. Concerning this matter, it is assumed that the stock price follows an inhomogeneous Poisson process, *i.e.* the number of jumps between t and $t+1$ is Poisson distributed:

$$P\left(\#\{\text{jumps}\}_t = k\right) = \exp\left(-\left|EED\right|_t\right)\frac{\left|EED\right|_t^k}{k!}, \quad k \in N_0, \tag{3.15}$$

where the intensity is equal to the total expected excess demand and each jump is linked to a price increase or decrease of 1 cent. In fact, this forms a way to transfer the agents' dynamics into price dynamics without any additional pararmeterization. Instead of the presented Poisson process, Lux and Marchesi (2000) use the following probabilities for a binary decision problem:

$$\pi_{\uparrow\ p} = max\left(0, \beta\left(ED + \mu\right)\right), \qquad \pi_{\downarrow\ p} = -min\left(0, \beta\left(ED + \mu\right)\right). \tag{3.16}$$

The main rationale for deviating from Lux and Marchesi (2000) is that otherwise additional calibration for β and the variance of μ would be necessary. Instead, the inhomogeneous Poisson process allows to determine both the expectation and the variance of the price jumps by only one factor, namely the EED. As the EED does not directly determine the price realization (jumps), the stochastic influence can be conserved, while the first two moments given by the EED represent different grades of financial market activity. Next, we turn to the derivatives market.

3.2.3 Derivatives Market

Heterogeneous Agents in a Dynamic System

Throughout this paper derivatives markets are reduced to European call options as this is sufficient to show how the agent-based model can work in derivatives markets. For the extension of the Lux (1998)-model to the derivatives market we also refer to three different types of agents: (1) derivatives long noise traders (dl), who expect to benefit from a long call position. (2) derivatives short noise traders (ds), who expect to benefit from a short call position. (3) derivatives fundamentalists (df), who want to make money by arbitrage. Furthermore, for the purpose of risk calculation we introduce three other variables: Firstly, let $frac$ denote the fraction of market participants who trade derivatives alongside equities. [21] Secondly, let T_{call} stand for the maturity of the contract (30 days).

[21] W.l.o.g. frac is set to 20 %. Note that h is always related to a certain underlying asset.

Finally, we introduce h as the 'systemic hedge ratio'. This indicates the fraction of all derivatives contracts used for the purpose of hedging. In other words these are covered call positions. We will analyze two possible cases $h = 0.25$ and $h = 0.75$. Based on this analysis, we argue that h represents a key ratio which should be monitored by the regulatory authorities in the future.
On the one hand, a derivatives long noise trader can be an optimist concerning the underlying asset. Hence he uses the derivative for additional speculation. On the other hand, a derivatives long noise trader can be a pessimist, when it comes to the underlying asset. In this case he uses the derivative to hedge his position. Thus the overall result for the first type of agent is given by

$$\#\{dl\}_t = \; frac\,(1-h)\,n_t^+ + frac\; h\; n_t^- . \tag{3.17}$$

This means that the number of derivatives long noise traders mainly depends on the number of optimists and pessimists in the underlying market (and thus, implicitly, on all dynamics involved there), as well as on the systemic hedge ratio assumed to be time-invariant.[22]
Similarly, a derivatives short noise trader can either be a pessimist or an optimist concerning the underlying asset. In the first case he uses the call option for additional speculation and aims at the option premium. In the second case, he hedges his position to the extent of the option premium. Thus the overall number of derivatives short noise traders is given by:

$$\#\{ds\}_t = \; frac\,(1-h)\,n_t^- + frac\; h\; n_t^+ . \tag{3.18}$$

Again, the number of agents can be traced back to the number of optimists and pessimists in the underlying market and to the systemic hedge ratio. From the dynamics of the other types of agents it is straightforward to obtain the number of derivatives fundamental traders:

$$\#\{df\}_t = frac\; N - (\#\{dl\}_t + \#\{ds\}_t) . \tag{3.19}$$

Note that the selected design is geared to the dimensions of options trading as described by Lakonishok, Lee, Pearson and Poteshman (2007) in their comprehensive exploration of investor behavior in the options market: There is a distinction made between long and short positions, while the purpose of investment - hedging or speculating - is also taken into account directly. Some of their results point towards trend-chasing, which, based on the majority opinion in the underlying market, is also integrated in our model.
Before turning to options prices derived from the agents' dynamics, it is necessary to explain how many contracts are included for the VaR computation in the

[22] A time-variant characterization is left for future research.

options market. [23] Without loss of generality and corresponding to the stock market model, this paper considers risk from only one asset in the derivatives market, i.e. from one call option. In the case of the stock, price paths for 250 trading days, in line with the Basel framework, have been simulated, while the call option matures after 30 days. That is why an arbitrary point in time has to be chosen, where the call starts to run.[24] It could be more advantageous to consider a revolving portfolio of options, but we decided to simplify as the presented design already foreshadows that the prevailing internal model methods neglect parts of the risk the agent-based model is able to capture. In particular, there are risks generated by a feedback mechanism described in the second next section.

Price Process

Having established the agents' dynamics in the derivatives market and given the transformation process in the underlying stock market, we can determine the expected excess demand in the option market. Here the EED of noise traders is defined by

$$EED_t^{n_c} = \left(\#\{dl\} - \#\{ds\}\right) vol_{model}. \tag{3.20}$$

Again fundamentalists start trading, if they identify arbitrage opportunities. Thus the expected excess demand generated by the derivatives fundamental traders looks similar to equation (3.13), whereby $c\,(t)$ stands for the call option market price:

$$EED_t^{f_c} = \frac{c_t^f - c\,(t)}{c\,(t)\,(1 + \hat{r}_M)} \#\{df\} vol_{model}. \tag{3.21}$$

As previously mentioned, the fundamental call price c_f is determined by the Black-Scholes formula. This is akin to the underlying stock market where the fundamental price was determined by the CAPM. Setting in $h = 0.5$ into equations (3.17) and (3.18) leads to an identical number of derivatives long and short noise traders. That is why, in this case, the call price is mainly driven by the Black-Scholes formula and depends on the market price of the agent-based model and other risk factors like implied volatilities and interest rates which are the same for all the models employed in the paper at hand. In order to emphasize the impact of monitoring the systemic hedge ratio, we focus on $h = 0.25$ and $h = 0.75$ which are arbitrary chosen but donot represent unrealistic cases.

[23] W.l.o.g. we assume an exchange ratio of 1:1 but this is independent of the portfolio one is interested in. One reason for the selected design is that we do not have to deal with the missing VaR subadditivity.

[24] In fact this explains the slightly different starting values. But note that the number of simulated price paths is the same - independent of which market one trades on.

Feedback

It is common knowledge that derivatives not only represent risky investments in their own market, but can also produce risk increasing feedback on the underlying market. Otherwise the fear of the so-called triple witching hour, where three kinds of derivatives simultaneously expire, would be unfounded. Regrettably, such spillover-effects are usually not considered in the standard models. We argue that it is possible to include the spill-over effect into the risk calculating model. At this juncture the systemic hedge ratio plays an important role.
In our model, feedback to the market for the underlying stock is determined by the final payoff $C(T) = \max(S - K, 0)$ with strike K and spot price S. Hence we differentiate two possible cases. First

$$(S-K)^{+} > 0. \tag{3.22}$$

From a risk management perspective, the most relevant feedback is cyclical. Equation (3.22) reveals a bullish market. Additional demand arises when the speculative part of derivatives short traders have to deliver the stock they do not own - similar to a situation arising from short-selling of the underlying stock. This leads to additional excess demand of

$$EED_{T_{call}}^{\text{call}\to\text{stock}} = \left(frac\,(1-h)\, n_{T_{call}^{-}}^{-} + \left(EED_{T_{call}^{-}}^{df} \right) \right). \tag{3.23}$$

In the second case the option will not be exercised, i.e.

$$(S-K)^{+} = 0. \tag{3.24}$$

Again, the most relevant feedback is cyclical. Equation (3.24) reveals a bearish market. Less demand arises when the speculative part of derivatives long traders have to sell stocks in order to rebalance their budget. Note the asymmetry of the payoff function. Thus losses only arise to the extent of the option premium. This leads to a lower excess demand [25] of

$$EED_{T_{call}}^{\text{call}\to\text{stock}} = -\frac{c\,(t=1)}{p\,(T_{call})} \left(frac\,(1-h)\, n_{T_{call}^{-}}^{+} + \left(EED_{T_{call}^{-}}^{df} \right) \right). \tag{3.25}$$

Summing up, the total expected excess demand of the stock in T_{call} is then given by

$$EED_{T_{call}} = EED_{T_{call}}^{n} + EED_{T_{call}}^{f} + EED_{T_{call}}^{\text{call}\to\text{stock}}. \tag{3.26}$$

Overall, this captures the fact that derivatives can amplify turbulence in the underlying market. The crucial point is: The lower the systemic hedge ratio,

[25] Under certain conditions - in a concrete manner, if the Black-Scholes formula anticipates the final payoff very well - the derivative fundamental (df) trader can lower the spill-over effect. But again this shows the limits of arbitrage, since it does not always work.

the higher the spill-over effect can be. As a consequence of the spill-over effect, the bubble frequency can increase. In that case, the spill-over effect plays exactly the role of the initial stochastic impulse as described in Section 3.2.1.

3.2.4 Currently Applied Internal Model Methods

Black-Scholes-Merton type Models

The models introduced by Black and Scholes (1973) and Merton (1973) assume perfect and complete capital markets. Specifically, this includes homogeneous expectations and rational behavior, no transaction costs [26], no short-selling constraints [27] and consequently no arbitrage. Moreover in its standard version, the underlying asset is assumed to follow the Geometric Brownian Motion, *i.e* .

$$dS\left(t\right)=\tilde{\mu}S\left(t\right)dt+\sigma S\left(t\right)dW\left(t\right), \qquad dW\left(t\right) \quad \sim N\left(0,dt\right). \tag{3.27}$$

By Ito's lemma we have

$$\ln\left(S\left(t\right)\right)=\ln\left(S_0\right)+\mu t+\sigma\sqrt{t}\epsilon, \quad \epsilon\sim N\left(0,1\right) \quad t>0, \qquad \tilde{\mu} \quad =\mu+\frac{\sigma^2}{2}. \tag{3.28}$$

The models by Black and Scholes (1973) and Merton (1973) also deliver (European) call and put prices. In the case of standard call options, and for $S=S\left(t\right)>0, 0\leq t<T$ we have

$$\begin{aligned} C\left(S,t\right)&=S\Phi\left(d_1\right)-K\exp\left(-r\left(T-t\right)\right)\Phi\left(d_2\right)\\ C\left(0,t\right)&=0, \quad C\left(S,t\right)\sim S \text{ for } S\to\infty, \quad C\left(S,T\right)=\left(S-K\right)^+\\ d_1&=\frac{\ln\left(S\ /K\right)+\left(r+\sigma^2/2\right)\left(T-t\right)}{\sigma\sqrt{T-t}}, \quad d_2=d_1-\sigma\sqrt{T-t}, \end{aligned} \tag{3.29}$$

In this paper we only deal with stock options[28] but the basic concept could also be applied to other derivatives - including interest and credit derivatives for which realizations of the risk factors are observable on a market. Hence, on a large scale, the call options can be taken as representative of derivatives market in general.

GARCH Models

In this section, we also deal with generalized autoregressive conditional heteroskedasticity (GARCH) models as an alternative approach to risk manage-

[26]The agent-based model (ABM) does not consider transaction costs directly. But, in the literature, one of the most discussed implications of existing transaction costs is that markets are no longer arbitrage-free. And this is certainly considered with the ABM.

[27]In the aftermath of the financial and the European sovereign-debt crises, several nations at least partly banned short selling in their financial markets.

[28]The market for put options can be modeled easily by simple changes of the corresponding models for call options.

ment, in order to cope with the issue of model uncertainty.[29] Although these models are less frequently employed in risk management than those of the Black-Scholes-Merton type, we discuss them in the section of the prevailing internal model methods. To keep the paper clearly, we consider GARCH only for the stock market.

Similar to the agent-based model we restrict ourselves to the stationary part in the description of the dynamics. Basically, the deterministic trend component is the same for all models, namely the one from Equation (3.28).[30]

GARCH(p, q) stock price returns are described by the following dynamics[31]:

$$sr_t = \sigma_t Z_t, \quad Z_t \overset{D}{\to} N(0,1), \tag{3.30}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i sr_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{3.31}$$

for a strictly positive process $(\sigma_t)_{t=1,2,\dots}$ and sr_t denoting stock returns. In order to specify the GARCH models, we follow a standard two-stage procedure. First, we estimate the best fitting AR model and carry out a Lagrange Multiplier (LM) test for ARCH in the residuals. As could be expected from Figure 3.6 in Appendix 3.5, we find strong ARCH effects for the sample 2008M1 - 2008M12, while for the sample 2007M1 - 2007M12 in most cases[32] the null hypothesis of no ARCH effects can only be rejected using a higher significance level. This result already indicates an important difference between ABM and GARCH, namely the fact that the latter can more or less only simulate the excess volatility that is already included to some extent in the calibration sample of the last 250 trading days.

Second, Equation (3.31) is specified starting from the order of ARCH found by the LM test. We use BIC information criterion in a general-to-specific manner to answer the questions of whether to reduce the order or to exclude GARCH effects (lags of σ_t) in addition to ARCH effects (lags of sr_t). In almost all cases we end up with GARCH(1,1) specifications with insignificant constant

[29] We include GARCH risk estimations, because in principle and as already mentioned in the introduction, this model type is able to address the phenomena of excess volatility and volatility clustering. Nevertheless, there are substantial differences in comparison with the agent-based results which will be discussed below.

[30] In rare cases, in which simulations with the drift from Equation (3.28) produce negative prices, we simplify for both the ABM and the GARCH simulations by reducing the (negative) drift to the extent that simulated prices are within the range of the BSM simulations. We leave it for future research to find a more sophisticated mechanism which precludes the emergence of negative prices categorically. However, this does not derogate the strands of argument in the paper at hand.

[31] Software packages usually allow for the estimation with the error Z_t being student-t-distributed, in particular with degree of freedom slightly below 30, in order to address 'fat tails' already at this point.

[32] In one case, A1EWWW, the null cannot be rejected based on sample 2007M1-2007M12. Compare the relatively low GARCH values-at-risk in Table 3.1. As the corresponding GARCH model must be misspecified to some extent, Appendix 3.3 presents results for asset 514000 instead.

and GARCH coefficients α_1, β_1 which in some cases add up close but never exactly to 1. Such GARCH(1,1) models are not uncommon in value-at-risk estimation, see for instance Gencay, Selcuk and Ulugülyagci (2003).

3.3 Calibration

3.3.1 Risk Factors in the Agent-Based Model

For the purpose of (overnight) risk calculation not too much data load should be necessary to compute the agent-based values-at-risk. Beside the stock prices, interest rates and volatilities from the last 250 trading days, which are also needed in the case of the currently applied internal model methods, there are dividend payout ratios, profits per share and trading volumes that have to be loaded for the agent-based computations. Appendix 3.3 summarizes parameters and risk factors.[33] In this context, note that model parameters in risk management models will usually be called risk factors, if their values are based on observable market data. As dividends are paid only once a year, the additional data is based on 5 years averages.

Dividends are equal to profits per share times the dividend payout ratio. They are necessary for both computations: i. the transition probabilities between noise and fundamental traders, ii. the fundamental price using the CAPM and Gordon's formula. From the latter, we can recursively determine the initially assumed growth rate of the company's cash flow.

The trading volume plays an important role in keeping simulated prices within a realistic range.[34] Instead of the direct calibration used by Lux and Marchesi (2000), we try to derive the trading volume of the model (vol_{mod}) from observable XETRA data (vol). Therefore we employ the following OLS regression

$$vol_{model} = \beta\ CAP\,(vol, N, p_1) + u, \qquad \text{where } CAP := \left(\frac{1}{vol}\right)^{\left(\frac{1}{p_1}\right)} N^{\left(\frac{1}{vol}\right)}. \tag{3.32}$$

Note that - besides the observable trading volume - the regressor CAP also considers the current market price and number of market participants in the model. The reasons for this form are as follows: The higher the real-world average trading volume (per transaction), the more difficult it is for the model to reproduce the volume given an initial price level and a fixed number of market participants.

[33] We encourage future research to robustify the parameter values, in particular those that are set time-constant, but an exploration of the full parameter space goes beyond the scope of this paper.

[34] On the one hand future risk management models should consider the systemic risk of bubbles. On the other hand risk management practice requires models which do not produce astronomical prices.

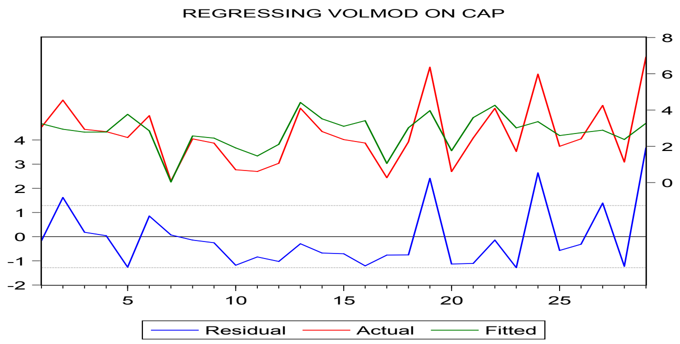


Figure 3.1: Actual and fitted model trading volumes (per transaction) for the DAX portfolio. Actual values are produced by the rule that existing bubbles should be kept within a realistic range, i.e. $p(t) \in [0, 2p(1)]\, \forall t$. The adjusted R^2 is equal to 0.46.

The higher the real-world price level, the lower real-world average trading volume tends to be.[35] The higher the number of participants in the model, the easier it is for the model to reproduce a higher trading volume. Of course, it is difficult to find an ultimate functional form for the interdependences mentioned above, but the presented one at least delivers an acceptable fit, see Figure 3.1. Finally, note that beside the fraction of market participants, who trade derivatives alongside equities, and the systemic hedge ratio no additional parameterization is necessary for the presented way of agent-based risk calculation in derivatives markets. If the above-mentioned quantities were monitored, this could be the basis to refine the corresponding parameter values. However, at the moment it is difficult to estimate exactly the fraction of options traders as long as a substantial part of the transactions is still running over-the-counter. Regarding the systemic hedge ratio, the values have been selected mainly for illustration purposes, namely to investigate the scenarios $h = 0.25$ and $h = 0.75$.

3.3.2 Risk Factors in other Models

In the BSM equations (3.27) and (3.28) $\tilde{\mu}$ denotes the drift parameter, μ the average of the empirical returns and σ the diffusion parameter or volatility. The latter two are taken from the observable market prices of the last 250 trading days. The only risk factor here is the stock itself. Its uncertainty is captured by the Wiener process W.

W.l.o.g., we choose an arbitrary option strike price K at the money.[36] Equation

[35] This statement is in particular true given a constant average budget.
[36] Note that the strike price is the same for all models employed in this paper.

(3.29) reveals the risk factors, which are the stock S (generated by the Wiener process), the interest rate r and the volatility σ. To pose interest rate risk and volatility risk we refer to historical values. The interest rate is taken from the one-month EURIBOR history and the volatility is conservatively approximated by the one-percent quantile of the VDAX NEW implicit volatility index which is also linked to an one-month maturity. The reason why implied volatilities are used is the dependence of the volatility on price developments and the strike price (volatility smile). For instance, implied volatilities of "out of the money" and "in the money" call options are generally higher than the implied volatility of "at the money" call options.
All GARCH coefficients in Equation (3.31) are estimated from observable stock price returns. As historical stock prices are not subject to revisions, the first sample 2007M1 - 2007M12 corresponds to a risk calculation at the end of 2007, while the second sample 2008M1 - 2008M12 assumes that the VaR computation has taken place at the end of 2008. This timing is the same for all models. In the GARCH simulations, the uncertainty is captured by the random process Z, while the starting value for the return volatility is identical to the sample standard deviations of the above-mentioned periods.

3.3.3 Simulation Effort and VaR Computation

For each method of risk calculation m price paths are taken into account, whereby each price path represents one trading year. Values-at-Risk (VaR) are calculated as the average 1% worst return that could arise, *i.e.* formally

$$\widehat{VaR} = \frac{1}{m}\sum_{i=1}^{m} \lceil q_{0.01}^{i} \rceil \left(sr_t\right), \quad P\left(sr_t \leq VaR\right) = 0.01, \tag{3.33}$$

where $q_{0.01}^{i}\left(sr_t\right)$ stands for the 1% quantile of the stock price returns of the i-th path, so that in 99% of all cases daily losses will not exceed the VaR.
In the case of the agent-based model, 40 elementary time units stand for one trading day because of the embedded intraday trading mechanism. Intraday price returns are cumulated before determining the 1% quantile so that the quantile, as with the other models, corresponds to the third worst of the 250 trading days. The whole simulation effort is then 5 times 250 times 40. Although a higher number of price paths leads to a visibly longer running time as in the case of the other methods, such an approach will still be practicable in terms of computability.[37]

[37] As a matter of simplification and for illustration purpose m is equal to 5. We also checked higher numbers without obtaining structurally different values-at-risk. Nevertheless, based on computational ressorces in banking practice a higher number should be required to robustify these results.

In the case of the Black-Scholes-Merton type model, an elementary time unit stands for one trading day. Here, the simulation effort is lower, namely 5 times 250. The same applies for the GARCH model. Note that it will not be the different definition of the elementary time unit that explains the significant differences in the VaR results, but the price path behavior, namely the fact that the agent-based model simulates excess volatility and that it does so disproportionately to the excess volatility, which could already be observed in the last 250 days' price history.

3.4 Simulation Results

3.4.1 Agent-based Simulations

Figures 3.2 - 3.5 present simulated price paths before and in the crisis for an arbitrary selected stock from the German stock index DAX. These paths could have been generated by the bank's internal risk management, if the agent-based method had already been required by the regulatory framework at this time. Basically, the ABM price paths and thus the corresponding VaR computation are closely related to a potential bubble or to the phenomena of excess volatility and volatility clustering. This is fully intended, since if these phenomena can occur in real-world trading indeed, their occurrence should also be simulated by the risk management.[38] Referring to DeGrauwe, Dewachter and Embrechts (1993), who were among the first to model exchange rates based on chaotic attractors, Lux (1998, p.145) explains the underlying mechanism of the agent-based financial market:

> 'The key mechanisms of [such] models are the following: (1) chartists' positive feedback reaction destabilizes the equilibrium in which price equals fundamental value, (2) an increasing strength of the fundamentalists' reaction upon differences between actual market price and fundamental value keeps in check the otherwise unstable oscillations.'

In other words, markets do not always tend to turbulences or 'non-efficient' states, but they do again and again. Precisely this point is reflected in the results of figures 3.2 - 3.5. In the simulations for the time before the crisis one turbulent market period can be found, in the simulations for the time during the crisis there are three.

As already mentioned, there is a similarity between ABM and GARCH volatility simulations. Compare Appendix 3.3, which also illustrates a change between

[38] At the moment those aspects are exclusively covered by stress tests, for which it is not clear, how much they affect the everyday business strategy.

financial markets periods of quiescence and periods of agitation. But resulting excess volatility periods of ABM and GARCH differ both in magnitude and in data reference. While the time-variation of the volatility in the GARCH depends directly on the sample for which the GARCH has been fitted, this is not clear for the chaotic data generating process of the agent-based model. However, we obtain for one (`A1EWWW`) of the three considered assets a similar pro-cyclical change of the price paths and corresponding values-at-risk when comparing ABM and GARCH results. The question is, if it is purely coincidental or irreproducible how the number of turbulent market phases changes in the ABM ? At first glance, one might think so, given the pro-cyclical effect for one of the assets. Certainly, work has to be done in order to completely disaggregate the individual effects in the relatively complex system of the agent-based simulations. On the other hand, there is an empirically verifiable reason why the intensity of turbulent market phases in the ABM behaves counter-cyclically. This reason is the trading volume neglected in the GARCH model. As a stylized fact the trading volumes of risky assets tend to increase with bull markets, while they tend to decline with bear markets or in times of crisis. In the model, as was discussed in Section 3.3, a lower trading volume has the general effect that the probability of price turbulences decreases by means of a lower excess demand which now can be compensated by arbitrageurs.[39]
In a nutshell, the increase in stock market value-at-risk (VaR) for stock `A1EWWW` from -11.40% (before the crisis) to -39.22% (in the crisis) is accompanied by the fact that the number of turbulent price periods within the simulation increases from sample to sample. Thus, for this asset we do not obtain a counter-cyclical development in contrast to the other two assets, for which VaR results are reported in Table 3.1.
Appendix 3.3 shows the results of agent-based risk calculation in the options market. As in the stock market, a substantial part of the risk generation arises from the interaction of the agents, especially when behavioral heuristics, such as the extrapolation of the actual trend or herding behavior, lead to price implosions and explosions. The fact that such turbulent market periods are also observable in real-world options market is well documented by Lakonishok et al. (2007). Subordinated risk generating factors are linked to the fundamental price process. Those are the aspects captured by the Black-Scholes formula.[40]

[39]Obviously, the decline of the average trading volume per transaction in the model from 2.9 to 2.0 for asset `A1EWWW` is not strong enough in contrast to the other two assets, for which VaR results are reported.
[40]Sometimes it is argued that there are less knocking outs observable on option than on stock markets. For some underlyings this might be true, for others not. However, exclusively using Black-Scholes formula leads to the fact that risk management models do not simulate the feedback effect between the markets.

Simulation before the Crisis for stock A1EWWW

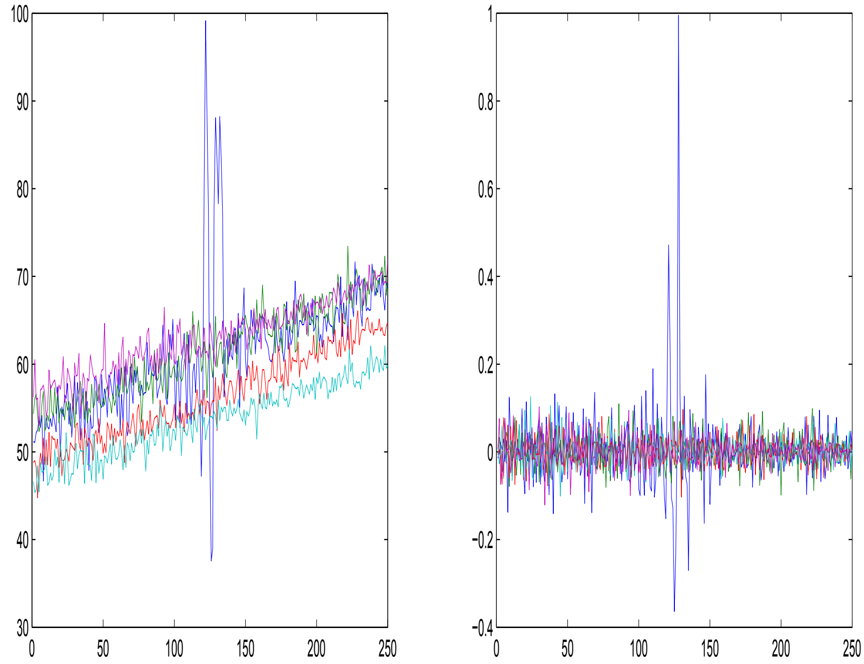


Figure 3.2: Simulated stock price paths (left-hand), the corresponding returns (right-hand), *i.e.* the relative change in prices. Prices are based on the risk factor sample *2007M1-2007M12*. One price path includes turbulent market periods. In financial literature, those are related to the concept of excess volatility. The ABM results in a Value at Risk equal to -11.40%.

Simulation before the Crisis for stock A1EWWW

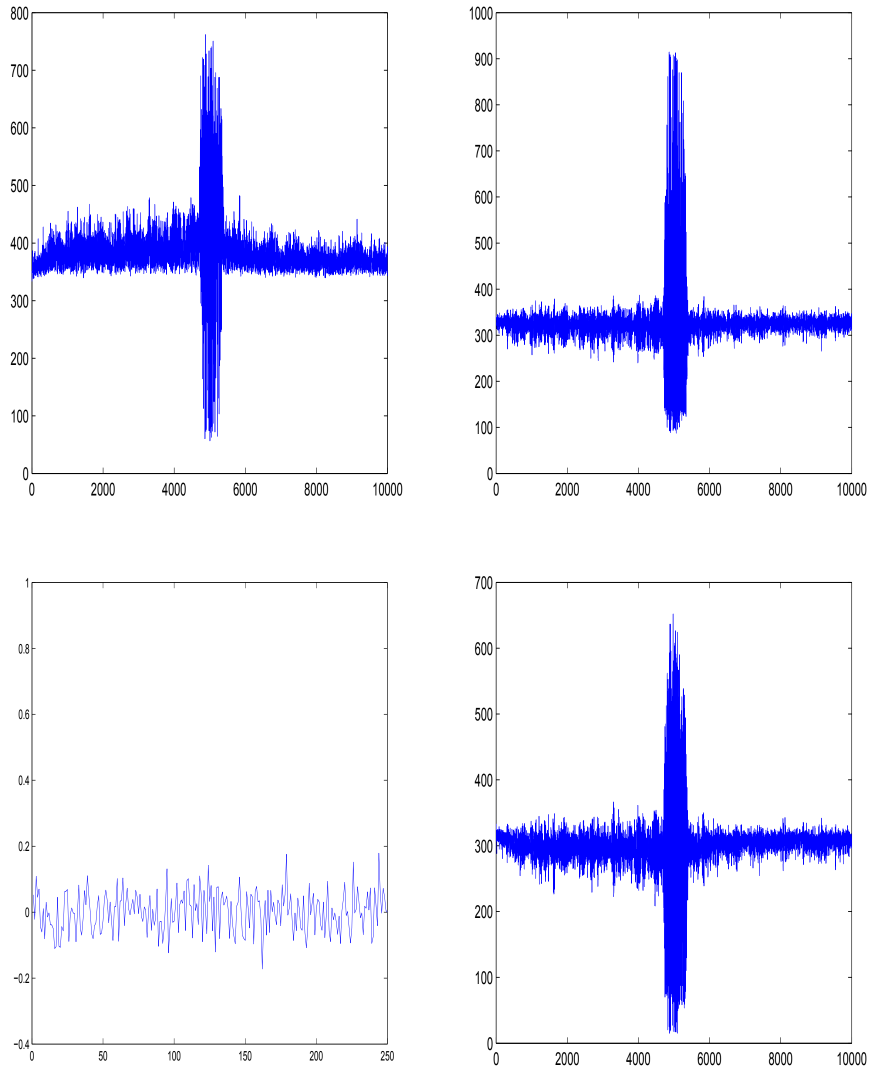


Figure 3.3: Fundamental price fluctuations (bottom left: scaling is the same as for the returns in Figure 3.2) and agents' dynamics for the price path with the highest volatility within the simulation. The top left figure plots the dynamics for the stock market fundamental trader, the top right for the optimistic noise trader and the bottom right for the pessimistic noise trader.

Simulation in the Crisis for stock A1EWWW

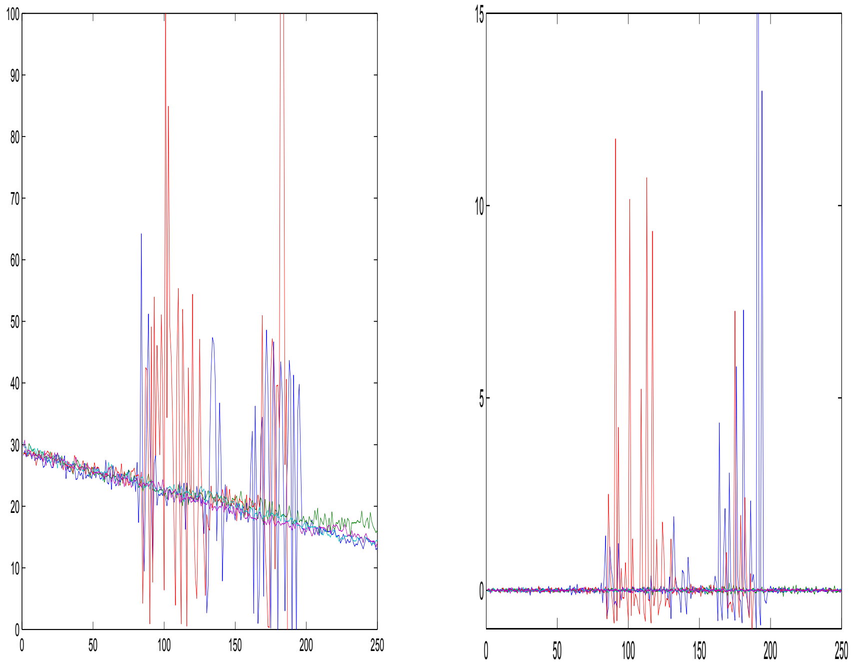


Figure 3.4: Simulated stock price paths (left-hand), the corresponding returns (right-hand), *i.e.* the relative change in prices. Prices are based on the risk factor sample *2008M1-2008M12*. Two price paths include turbulent market periods. In financial literature, those are related to the concept of excess volatility. The ABM results in a Value at Risk equal to -39.22%.

Simulation in the Crisis for stock A1EWWW

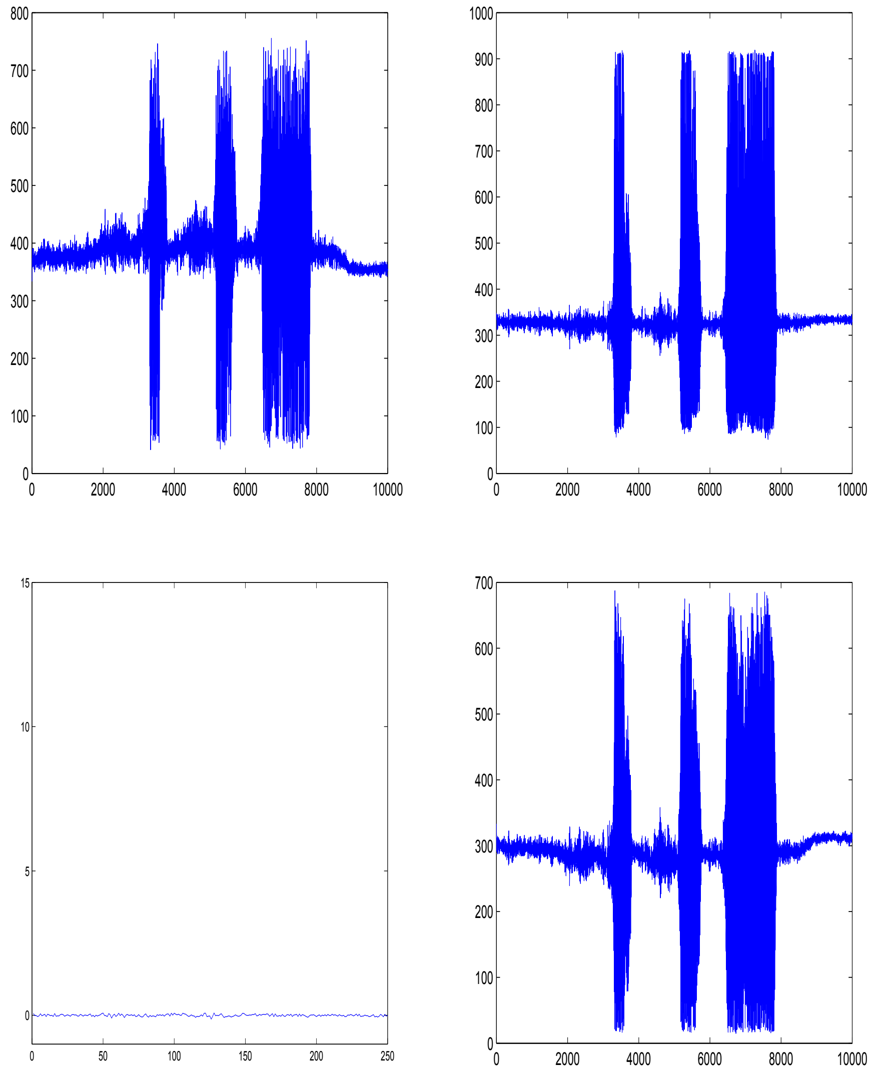


Figure 3.5: Fundamental price fluctuations (bottom left: scaling is the same as for the returns in Figure 3.4) and agents' dynamics for the price path with the highest volatility within the simulation. The top left figure plots the dynamics for the stock market fundamental trader, the top right for the optimistic noise trader and the bottom right for the pessimistic noise trader.

For the call position on the arbitrary chosen asset the VaR essentially retains its level: -39.52% before the crisis and -45.48% during the crisis. Given a certain degree of simulation variation, such numbers do not describe pro-cyclical dynamics. When looking at the different price paths, there is also a parallel line to papers that draw conclusions about the validity of the efficient-market hypothesis, at least in terms of IIDness of the increments.[41] Not all, but several price paths produce excess volatility. This fits the empirical finding that a rejection of the efficient-market hypothesis is strongly sample dependent (Malkiel, 2003, for a survey). Hence, some stylized facts, compare Appendix 3.5, are better reproduced by the ABM price paths than by the BSM paths which assume a general validity of the efficient-market hypothesis.

3.4.2 Model Comparison

Price Path Comparison

Appendix 3.3 displays the simulated Black-Scholes-Merton type stock price paths. In there, Figure 3.11 illustrates the results for a point in time before the crisis and Figure 3.12 for a point in time in the crisis. Appendix 3.3 presents the corresponding results for the derivatives market.[42] The average stock price paths for the different points in time reflect bearish and bullish markets. In the stock market, the VaR for the selected asset changes from -2.88% (before the crisis) to -7.42% (during the crisis) which is representative for a typical pro-cyclical BSM VaR behavior. In the derivatives market, the VaR also increasese from -28.65% (before the crisis) to -61.79% (during the crisis). Note that the higher level of the VaR in derivatives market is to some extent connected to relatively small prices at the time of the call option maturity.

The increments of the Wiener process are stationary and normally distributed with mean zero and variance $t - s$ for two points in time s, t, compare Equation (3.28). As a result, the range of simulated BSM prices increases with the simulation horizon. In particular, the range of prices at the end of the simulation horizon is wider with the BSM as with the ABM. However, the variance between t and $t + 1$ does not change for all times. Since the p% worst daily loss is ultimately the crucial factor for the value-at-risk computation, the constant variance over time derogates the appropriateness of the BSM for risk management purposes.

The differences between ABM and GARCH price paths have already been discussed. In the stock market (Appendix 3.3), the GARCH VaR for the

[41] Given an underlying IID information process, several studies starting with Lo and MacKinlay (1988) have convincingly rejected a stock price random walk based on econometric properties like predictability through significant serial correlation.

[42] MATLAB code for all models is available on request so that risk calculation for all the stocks in the index and the samples 2007M1 - M12, 2008M1 - M12 can be executed.

VaR market /sample	Asset	ABM		BSM		GARCH		Maximum
stock market	A1EWWW	-11.40%		-2.88%		-1.60%		ABM
2007M1-M12	514000	-32.52%		-3.37%		-5.69%		ABM
	766403	-11.81%		-4.09%		-4.45%		ABM
stock market	A1EWWW	-39.22%	↑	-7.42%	↑	-3.41%	↑	ABM
2008M1-M12	514000	-38.23%	→	-10.52%	↑	-11.44%	↑	ABM
	766403	-9.22%	→	-9.08%	↑	-14.10%	↑	GARCH
derivatives market	Call on A1EWWW	-39.52%		-28.65%		-		ABM
2007M1-M12	Call on 514000	-64.13%		-32.08%		-		ABM
	Call on 766403	-36,35%		-59.34%		-		BSM
derivatives market	Call on A1EWWW	-45.48%	→	-61.79%	↑	-		BSM
2008M1-M12	Call on 514000	-44.81%	→	-68.86%	↑	-		BSM
	Call on 766403	-7.53%	↓	-43.51%	→	-		BSM

Table 3.1: Values-at-risk and the cyclical properties of different samples, markets and assets with respect to the comparison of the prevailing internal model methods (BSM and GARCH) and the agent-based model (ABM). Samples describe the market periods based on which risk factors are calibrated. In general, the ABM values-at-risk values turn out to be higher and more countercyclical than the other ones. The column asset refers to the German Securities Number. The arrows plot the change of the VaR for each asset in the portfolio from the time before to the time after the financial crisis had started. $\rightarrow$ stands for a relative change of the values-at-risk within the interval $[-50\%, +50\%]$, while $\uparrow, \downarrow$ denote stronger changes. Overall these findings are the basis for recommending the integration of agent-based models into the regulatory framework.

selected asset changes from -1.60% (before the crisis) to -3.41% (during the crisis). In general, also GARCH values-at-risk tend to be pro-cyclical.

Value-at-Risk Comparison

Table 3.1 summarizes the values-at-risk of various asset for the different calibration periods and for all models. Calibration period 2007M1-M12 is before the crisis and period 2008M1-M12 in the crisis so that we can draw conclusions about the cyclicality of values-at-risk. The evolution of the VaR of a certain asset is indicated by the arrows in Table 3.1. The VaR point estimates in the paper at hand cannot be unique, so we use a broad tolerance limit, $+/-50\%$ of the initial value, to avoid misinterpretation of the changes in VaR. In contrast to the graphical analysis of the previous section, table 3.1 considers three assets for both the stock and the derivatives market in order to base conclusion about the VaR level on a solid foundation.

The results can be interpreted as follows: First, ABM VaRs are within a realistic range[43], while the absolute level or the amount of change of a single value may be questioned. Second, the values-at-risk (VaR) of the models differ significantly, which means that the regulatory framework should be extended due to model uncertainty.[44] Third, BSM and GARCH VaRs for the stock market are below eight percent[45] in pre-crisis time, while they are at or above eight percent in the crisis. In contrast, all ABM VaRs are at or above eight percent. In general, values above a certain threshold, *e.g.* eight percent, can be considered as capital requirements, which serve to map the systemic risk (of turbulent market periods) to the individual financial instruments apart from their intrinsic market price risk. Fourth, ABM VaRs tend to behave less pro-cyclically due to the decreasing trading volumes in the crisis.

3.4.3 VaR, Excess Volatility and Market Stability

This section develops the implications of the previous values-at-risk results for a regulator, while the implementation of regulatory measures is concretized in the next section. One of the main objectives of a regulator is the reduction of excess volatility and, consequently, the preservation of financial market stability. While abstracting from other objectives here, the regulator faces the question whether (higher) capital requirements can reduce excess volatility. To establish this link we review some evidence from the literature as the models at hand serve to compute the ex-ante value-at-risk and do not investigate how return volatility is affected ex-post by (higher) value-at-risk-based capital requirements.

Buss, Dumas, Uppal and Vilkov (2013) explore the effects of different regulatory measures against the background of reducing the volatility of stock market returns. They employ a dynamic stochastic general equilibrium (DSGE) model of a production economy with heterogeneous investor beliefs, while the investors can trade riskless and risky assets in financial markets. The authors conclude that 'only the leverage constraint is effective in reducing stock-market volatility, and this is accompanied by positive effects on the real sector: an increase in the levels of consumption growth and investment growth, and a decrease in their volatilities' (Buss et al., 2013, p.1). As their leverage constraint is not based on

[43] All ABM VaRs are considerably below 100% and in some cases below the values derived from the other models.

[44] This statement refers to the disputable existence of a *true* data generating process which can be reproduced by a single model.

[45] The regulatory standard approach required an eight percent and in reaction to the crisis now requires a ten point five capital requirement of risk-weighted assets, while in our calculations, the current price (the nominal) represents the reference value.

a value-at-risk and the banking sector not explicitly modeled, their impulse response cannot be used to quantify the effects of higher ABM VaRs. However, we can argue qualitatively using two channels of the bank's balance sheet. *Ceteris paribus* , a higher VaR limits the business operations of the bank in two ways. In a direct manner, banks as investors can only invest less in the risky asset (less proprietary trading). Indirectly, a higher VaR reduces the credit supply of a single bank and limits the amount of borrowing for their clients and can therefore work as a leverage constraint.
Hellmann, Murdock and Stiglitz (2000) analyze the effects of capital requirements from a micro perspective. While they find a short-term incentive to reduce investment in the risky asset, they also mention a future conflictive effect linked to the banks' franchise values. However, their conclusion is not against capital requirements, but to supplement them by deposit-rate ceilings or asset-class restrictions. Repullo (2004) additionally states: 'Building on Hellmann et al. (2000), we have shown that for a particular model of imperfect competition in the deposit market, both instruments [capital requirements and deposit rate ceilings as regulatory tools] are in general effective in preventing the banks from taking excessive risks' (Repullo, 2004, p.25). His best performing solutions are 'risk-based capital requirements'. This characterization also includes agent-based values-at-risk.

3.4.4 Policy Implications

Capital Requirements
Summarizing the values-at-risk results from the previous sections, two issues are striking. Firstly, on the stock market the ABM VaR tends to be higher than the BSM VaR and the GARCH VaR. This points to a tightening of capital requirements in line with the response of regulatory authorities to the recent financial crisis. Secondly, the ABM-values-at-risk turn out to be more counter-cyclical than the BSM VaR and the GARCH VaR. Hence, including the agent-based model into the regulatory framework would provide a means of quantifying the neglected elements of systemic risk, and the resulting capital requirements would be shaped in a more counter-cyclical manner. If such a risk assessment is installed, systemic risk will be measured directly in the most developed parts of the Basel framework. These are the risk management models linked to the first pillar of Basel II.

Monitoring Derivative Markets

In contrast to the standard hedge ratio determining the contracts that are necessary to safeguard an individual portfolio amount against losses from price fluctuations, the systemic hedge ratio stands for the fraction of all derivative contracts in the market used for the purpose of hedging. In this section, we illustrate why this ratio should be monitored by regulatory authorities through setting a minimum threshold.

Table 3.2 shows the values-at-risk for different systemic hedge ratios. For the majority of assets it is hard to find significant feedback from the derivatives market on the stock market. This should not be too surprising, since in the model other effects can dominate. Indeed spill-over effects do not always occur. But in two cases such an effect can be isolated. Here, it is $EED_{T_{call}}^{\text{call}\rightarrow\text{stock}}$ being significantly different from zero and reflecting the additional demand from expiring derivatives contracts. Following the rationale from Section 3.2.1, the feedback effect can replace the first stochastic impulse from outside the model and thus create additional price fluctuations (potentially higher values-at-risk) in the underlying market. Compared to such a situation, a higher number of market participants using the derivative for the purpose of hedging instead of speculation (a higher systemic hedge ratio) can significantly lower the value-at-risk, as the corresponding $EED_{T_{call}}^{\text{call}\rightarrow\text{stock}}$ decreases.

One means of (institutional) implementation of a minimum systemic hedge ratio can be the shifting of OTC transactions into the regulated market (stock exchange), which corresponds to the intention to create centralized Designated Clearing Organizations (DCOs), see Chappe and Semmler (2012). Details for implementation go beyond the scope of this paper, but the necessary condition for hedging is the possession of the underlying asset so that in case of a systemic hedge ratio below the target threshold speculative derivatives transaction (transaction without having the underlying asset in the portfolio) could be prohibited.

At the end of this section, it is worth mentioning that in the literature also the occurrence of rational bubbles is discussed. This is based on Tirole (1985), who argues 'that bubbles are not inconsistent with optimizing behavior and general equilibrium'. In contrast, the presented model will be classified as one that deals with the so-called irrational bubbles. However, this distinction does not change the fact that somebody has to bear the costs at the moment where bubbles burst in order to guide the economy towards its previous state. In this way, the present paper opts for making the financial system internalize the external costs of a bubble.

VaR market/sample	Asset	ABM $h = 0.25$	ABM $h = 0.75$	Change
stock market	`A1EWWW`	-11.40%	-15.00%	$\rightarrow$
2007M1 - M12	514000	-32.52%	-16.28%	$\downarrow$
	766403	-11.81%	-14.68%	$\rightarrow$
stock market	`A1EWWW`	-39.22%	-14.49%	$\downarrow$
2008M1 - M12	514000	-38.23%	-39.01%	$\rightarrow$
	766403	-9.22%	-8.62%	$\rightarrow$

Table 3.2: Values-at-risk for different systemic hedge ratios. The arrows plot the corresponding change of the VaR for each asset. $\rightarrow$ stands for a relative change of the values-at-risk within the interval $[-50\%, +50\%]$, while $\uparrow, \downarrow$ denote stronger changes. For the majority of the assets it is hard to find significant feedback from the derivatives market on the stock market, but in two cases such an effect can be isolated.

3.5 Conclusion

From a micro perspective the paper at hand recommends the internalization of the external costs of bubbles. A comprehensive regulatory approach (Basel IV), including agent-based (ABM) risk calculation, could improve the steering effect of capital requirements as we find reasonable values-at-risk through the ABM which differ from existing models both in their level as well as in their cyclical properties. Such capital requirements should be demanded of all major market players, including non-bank actors like hedge funds.
The presented model also demonstrates the need of regulatory authorities to find out how many derivatives per underlying asset are used market-wide for the purpose of hedging (systemic hedge ratio). One way of (institutional) implementation could be the obligation to employ stock exchanges also as trading platforms for Over-The-Counter products. The agent-based model suggests to monitor the systemic hedge ratio by means of a minimum level as a lower value reduces the risk of feedback effects generating excess volatility in the underlying market. Agent-based risk management merits further research with respect to an even more granular calibration and and leaves room for the extension to other types of financial instruments which are not covered in this paper.

Appendix B

B.1 Transition Probabilities

This section explicitly lists the transition probabilities from Equation (3.3):

$$\pi^t_{+-} = \min\left(v_1 \frac{n^-_t}{N} \exp\left(-\left(a_1 \frac{n^+_t - n^-_t}{n^n_t} + a_2 \frac{p'(t)}{p(t)}\right)\right), 1\right)$$

$$\pi^t_{+f} = \min\left(v_2 \frac{n^f_t}{N} \exp\left(-a^{t,e}_3 \left(\frac{d_t + \frac{1}{v_2} p'(t)}{p(t)} - R - \left|\frac{p^f_t - p(t)}{p(t)(1+\hat{r}_M)}\right|\right)\right), 1 - \pi_{+-}\right)$$

$$\pi^t_{-+} = \min\left(v_1 \frac{n^+_t}{N} \exp\left(a_1 \frac{n^+_t - n^-_t}{n^n_t} + a_2 \frac{p'(t)}{p(t)}\right), 1\right)$$

$$\pi^t_{-f} = \min\left(v_2 \frac{n^f_t}{N} \exp\left(-a^{t,e}_3 \left(R - \frac{d_t + \frac{1}{v_2} p'(t)}{p(t)} - \left|\frac{p^f_t - p(t)}{p(t)(1+\hat{r}_M)}\right|\right)\right), 1 - \pi_{-+}\right)$$

$$\pi^t_{f+} = \min\left(v_2 \frac{n^+_t}{N} \exp\left(a^{t,e}_3 \left(\frac{d_t + \frac{1}{v_2} p'(t)}{p(t)} - R - \left|\frac{p^f_t - p(t)}{p(t)(1+\hat{r}_M)}\right|\right)\right), 1\right)$$

$$\pi^t_{f-} = \min\left(v_2 \frac{n^-_t}{N} \exp\left(a^{t,e}_3 \left(R - \frac{d_t + \frac{1}{v_2} p'(t)}{p(t)} - \left|\frac{p^f_t - p(t)}{p(t)(1+\hat{r}_M)}\right|\right)\right), 1 - \pi_{f+}\right).$$

Note that the sign of the ingredient of the exponential function ensures that if a certain transition probability is high (*e.g.* π_{+-}), its inverse (*e.g.* π_{-+}) will be low.

B.2 Anomalies or Stylized Facts ?

Figure 3.6 presents the stock price returns of two different samples - one before and one during the recent financial crisis. Both are part of the presented risk calculation, where for our findings the exact temporal delimitation is of secondary importance. What immediately attracts the attention in the sub-figure on the left are the peaks of the new economy and of the last crisis' bubble [46].

[46] Of course both bubbles differ in their cause and their effects. Nevertheless, in light of their absolute levels they seem to be comparable and, although the paper at hand concentrates on the financial crisis, we wanted to highlight the new economy experience to emphasize the fact that bubbles are recurring phenomena.

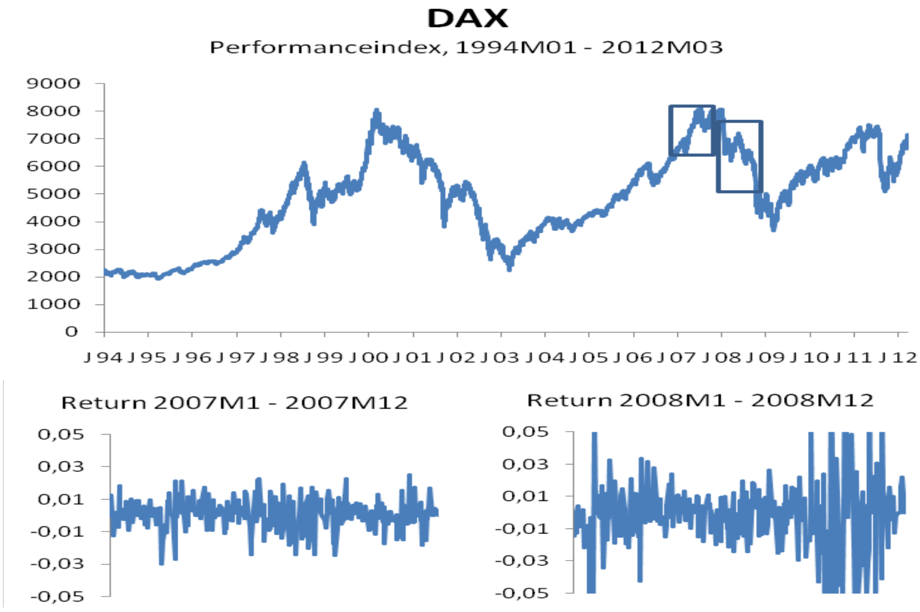


Figure 3.6: Stock prices and stock price returns for different samples reflecting the phenomena of excess volatility and volatility clustering - both often linked to bubbles.

Illustrated on the right we find periods where weak market-efficiency could hold (lower left) and where not (lower right). Additionally, excess volatility is only striking at the beginning and at the end of the lower right sample, interrupted by a relatively calm market period. The latter is due to Lehman Brothers declaring bankruptcy in September 2008. Summarizing, one can speak of volatility cluster. In this context, it is remarkable how matter-of-factly much of the literature used the terms anomaly or curiosa to describe the above-mentioned empirical facts - as if market-efficiency and no-arbitrage phases were the only natural state of financial markets. In contrast, Pagan (1996) in his seminal contribution paints a complete picture where (for univariate financial time series) he examines the questions of (1) stationarity, (2) independent distributions over time (which is linked to non-linear modeling and volatility cluster), (3) the existence of moments and of (4) normally distributed returns (which are both linked to excess volatility). The permanent change between calm and turbulent market periods also fits the analysis of Chen et al. (2001, page 1) who find that 'explicit tests for non-linearity and dependence also give very unstable results in that both acceptance and strong rejection of IIDness can be found in different realizations.' Finally Lakonishok et al. (2007) ascertain that, at least for the less sophisticated option traders, a special trading behavior can be observed in bubble times.

B.3 Parameters and Risk Factors

Agent-based model (ABM) - stock market transition process			
N	number of market participants	set time constant	1000 w.l.o.g Affects the simulation run time.
v_1, v_2	revaluation frequency	set time constant	$0.6, 0.4$ For a motivation see Section 3.2.1.
a_1, a_2 a_3	weighting factor of behavioral heuristics	set time constant or stochastic	$0.5, 0.5, \sim N\left(\mu_3^a, (U\,[0, 0.01])^2\right)$ For a motivation see Section 3.2.1.
Stock market: ABM fundamental price, GARCH and Black-Scholes-Merton price process			
b	dividends payout ratio	sample average (5 trading years)	www.ariva.de
d_t	dividends	sample average (5 trading years)	www.ariva.de
$\tilde{\mu}$	drift	sample average (250 trading days)	www.ariva.de
σ_p^2	diffusion	sample variance (250 trading days)	www.ariva.de
α, β	GARCH coefficients	estimated from daily returns	www.ariva.de
Agent-based model - market price processes			
vol	trading volume	sample average (5 trading years)	www.ariva.de
vol_m	model trading volume	estimated by OLS	For a motivation see Section 3.3
Agent-based model (ABM) - derivatives market transition process			
$frac$	market participants trading derivatives	set time constant	200 w.l.o.g. Affects the simulation run time.
h	systemic hedge ratio	set time constant	0.25 or 0.75 For a motivation see Section 3.2.3
Derivatives market: ABM fundamental price and Black-Scholes-Merton price process			
r	interest rates	sample average (250 trading days)	1M EURIBOR
σ^2	implied volatilites	sample average (250 trading days)	VDAX NEW index

Table 3.3: Model parameters and risk factors. Samples for risk factors before (2007M1 - 2007M12) and in (2008M1 - 2008M12) the crisis consist of 250 trading days due to the BASEL regulatories.

B.4 Agent-Based Simulations - Derivatives Market

Simulation before the Crisis for a call on A1EWWW

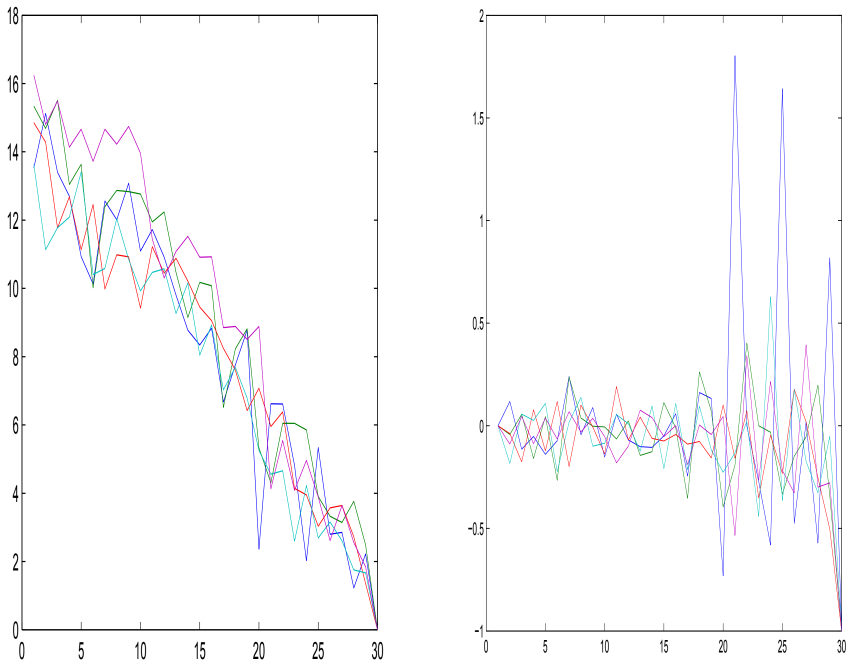


Figure 3.7: Simulated option price paths (left-hand), the corresponding returns (right-hand). Prices are based on the sample *2007M1-2007M12* as well as on a 30 days maturity. Some of the ABM simulations exhibit higher price fluctuations. In financial literature, those are related to the concept of excess volatility. The ABM results in a Value at Risk equal to -39.52%.

Simulation before the Crisis for a call on A1EWWW

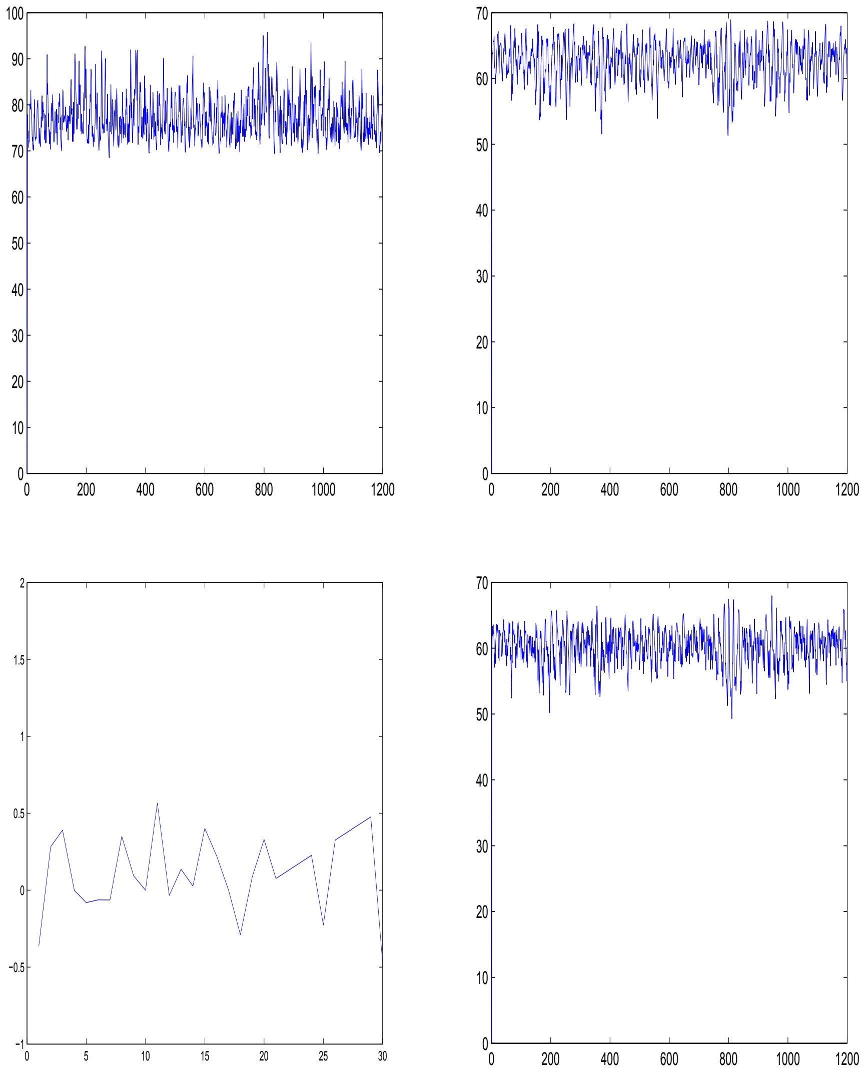


Figure 3.8: Fundamental price fluctuations (bottom left: scaling from the returns in Figure 3.7) and agents' dynamics for the price path with the highest volatility within the simulation. Top left: derivatives market fundamentalist, top right: derivatives long noise trader and bottom right: derivatives short noise trader. In contrast to other simulation runs, the trader classes remain occupied relatively stable over time.

Simulation in the Crisis for a call on A1EWWW

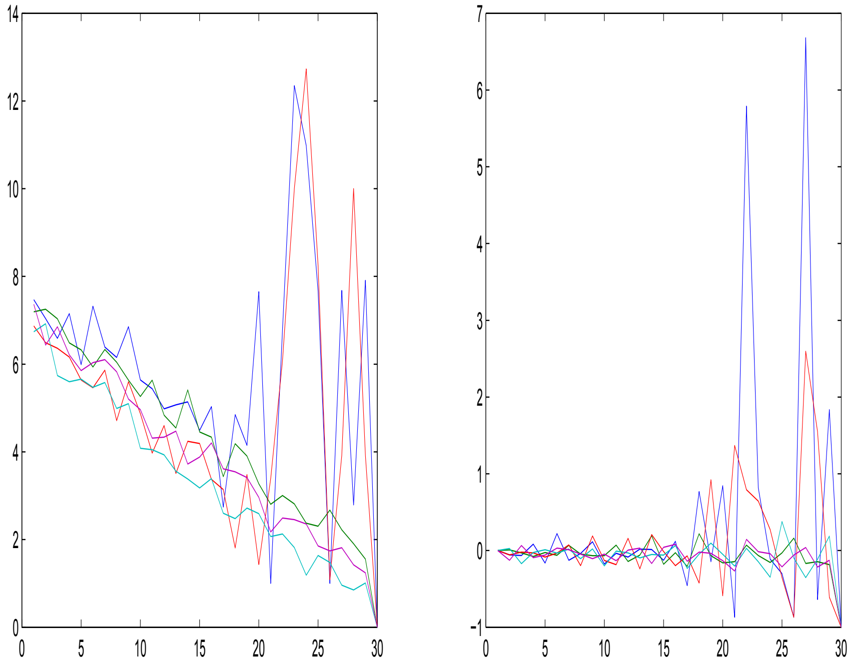


Figure 3.9: Simulated option price paths (left-hand), the corresponding returns (right-hand). Prices are based on the sample *2008M1-2008M12* as well as on a 30 days maturity. Some of the ABM simulations exhibit higher price fluctuations. In financial literature, those are related to the concept of excess volatility. The ABM results in a Value at Risk equal to -45.48%.

Simulation in the Crisis for a call on A1EWWW

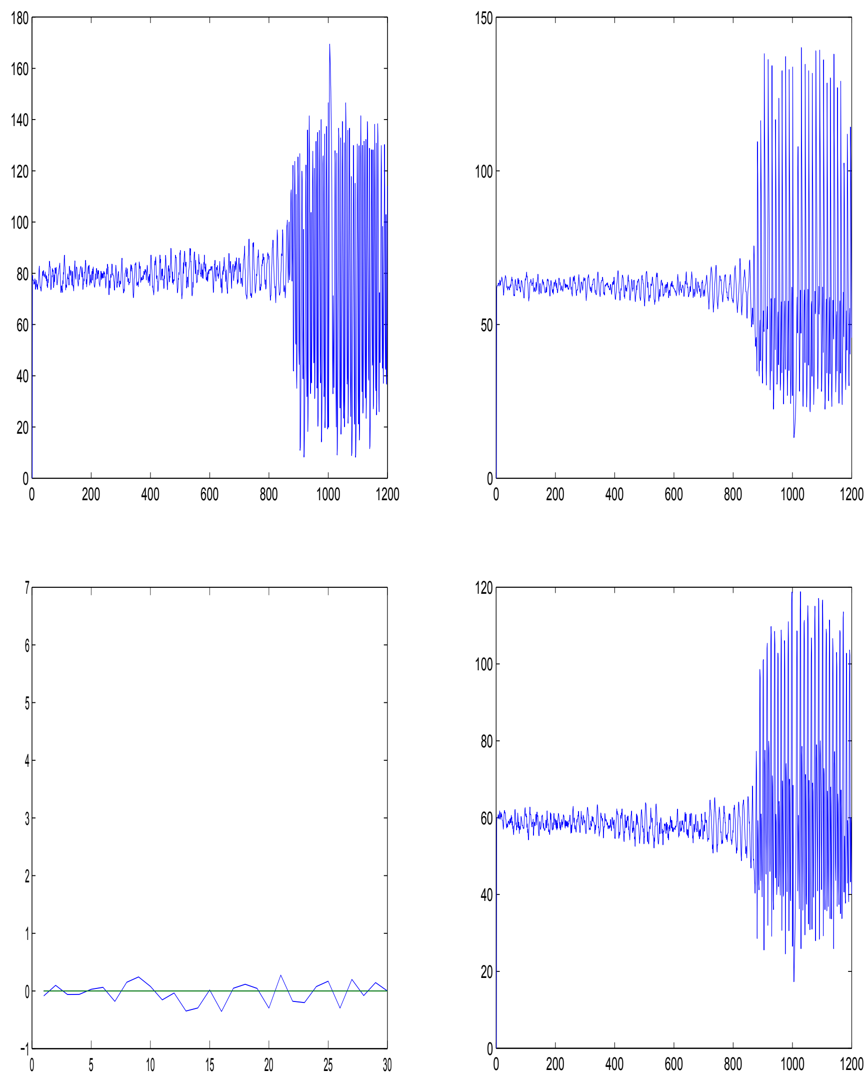


Figure 3.10: Fundamental price fluctuations (bottom left: scaling from the returns in Figure 3.9) and agents' dynamics for the price path with the highest volatility within the simulation. The top left figure plots the dynamics for the derivatives market fundamentalist, the top right for the derivatives long noise trader and the bottom right for the derivatives short noise trader.

B.5 Black-Scholes-Merton type Simulations - Stock Market

Simulation before the Crisis for stock A1EWWW

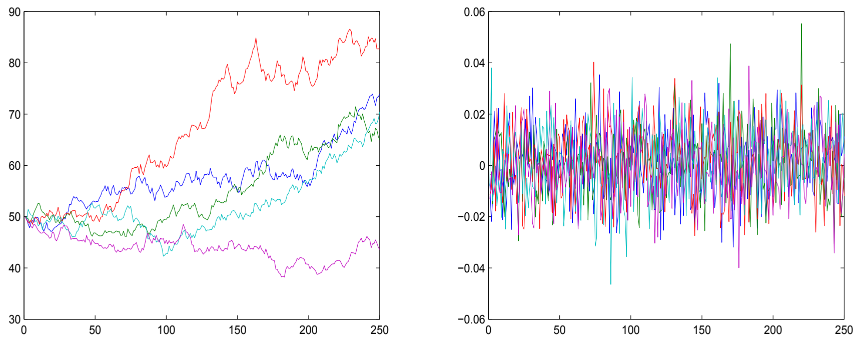


Figure 3.11: Simulated price paths (left-hand) and the corresponding returns (right-hand) with a Black-Scholes-Merton(BSM) type simulation, where drift (0.0011) and diffusion (0.0131) parameters are based on the sample *2007M1-2007M12*. The BSM results in a Value at Risk equal to -2.88%.

Simulation in the Crisis for stock A1EWWW

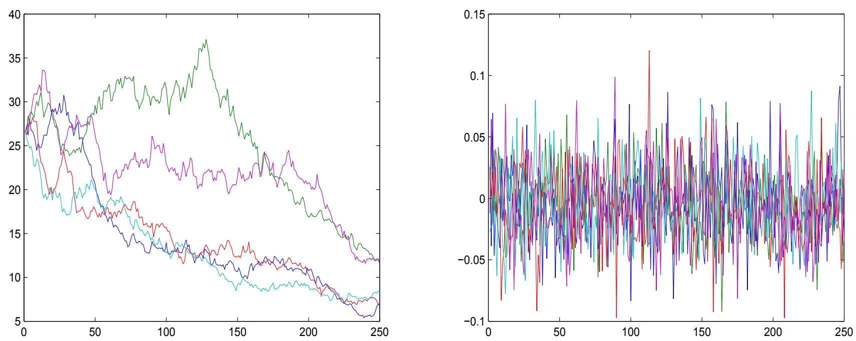


Figure 3.12: Simulated price paths (left-hand) and the corresponding returns (right-hand) with a Black-Scholes-Merton(BSM) type simulation, where drift (-0.0025) and diffusion (0.0304) parameters are based on the sample *2008M1-2008M12*. The BSM results in a Value at Risk equal to -7.42%..

B.6 Black-Scholes type Simulations - Derivatives Market

Simulation before the Crisis for a call on A1EWWW

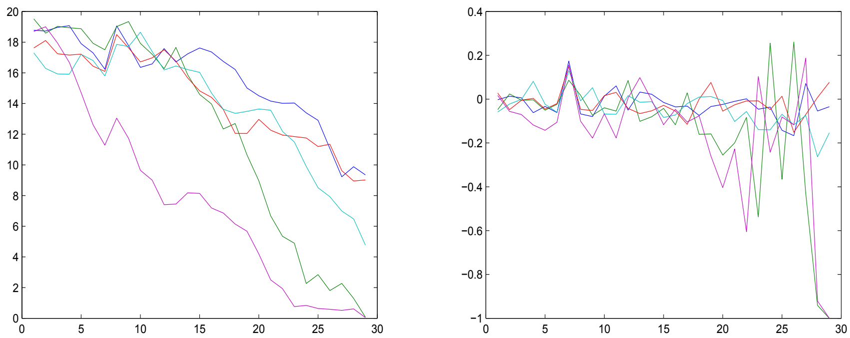


Figure 3.13: Simulated option price paths (left-hand) and the corresponding returns (right-hand) with BSM. The BSM uses the Black-Scholes formula, where volatility $\in [0.134, 0.314]$, interest rates $\in [0.0417, 0.0495]$ and stock prices are based on the sample *2007M1-2007M12*. This results in a Value at Risk equal to -28.65%.

Simulation in the Crisis for a call on A1EWWW

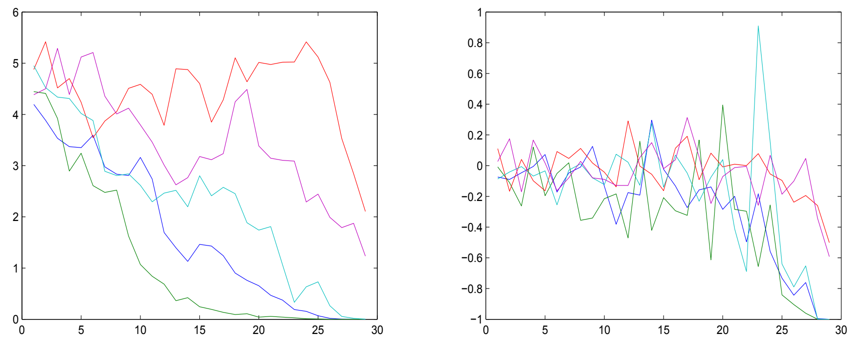


Figure 3.14: Simulated option price paths (left-hand) and the corresponding returns (right-hand) with BSM. The BSM uses the Black-Scholes formula, where volatility $\in [0.173, 0.832]$, interest rates $\in [0.0260, 0.0361]$ and stock prices are based on the sample *2008M1-2008M12*. This results in a Value at Risk equal to -61.79%.

B.7 GARCH Simulations - Stock Market

Simulation before the Crisis for stock 514000

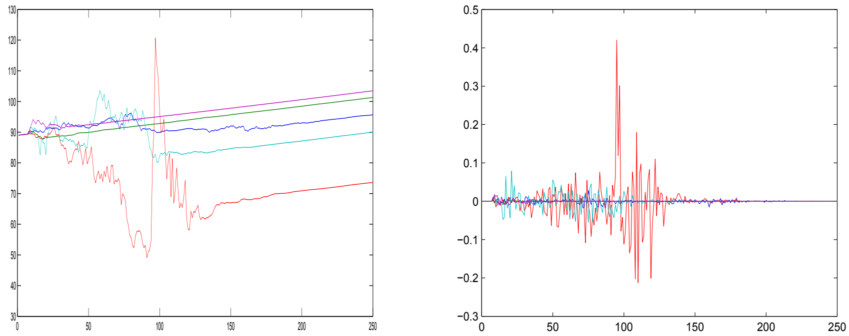


Figure 3.15: Simulated price paths (left-hand) and the corresponding returns (right-hand) with a GARCH model, where coefficients, initial prices and volatility are based on the sample *2007M1-2007M12*. The GARCH results in a Value at Risk equal to -5.69%.

Simulation in the Crisis for stock 514000

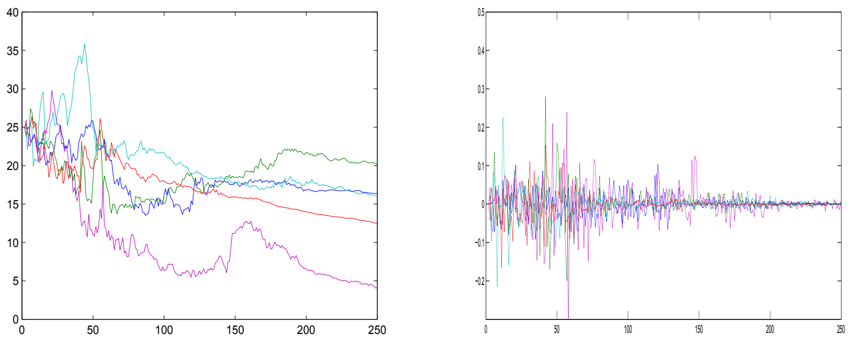


Figure 3.16: Simulated price paths (left-hand) and the corresponding returns (right-hand) with a GARCH model, where coefficients, initial prices and volatility are based on the sample *2008M1-2008M12*. The GARCH results in a Value at Risk equal to -11.44%.

Bibliography

Alicke, M. D. and Govorun, O. (2005), The Better-Than-Average Effect, *in* 'The Self in Social Judgment', M. D. Alicke, J. I. Krueger and D. A. Dunning (Eds.), pp. 86–106. Psychology Press.

Anas, J., M. Billio, L. F. and Mazzi, G. (2008), 'A system for dating and detecting turning points in the euro area', *The Manchester School* **76**, 549–577.

Ang, A., Piazzesi, M. and Wei, M. (2006), 'What does the yield curve tell us about GDP growth ?', *Journal of Econometrics* **131**, 359–403.

Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R. and Tayler, P. (1996), Asset Pricing Under Endogenous Expectations in an Artificial Stock Market, SSRN Discussion Paper 2252, Social Science Research Network.

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996), 'Fractionally integrated generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **74**, 3–30.

Basel Committee on Banking Supervision (2006), International convergence of capital measurement and capital standards: A revised framework, Technical report, Bank For International Settlements. Comprehensive Version.

Basel Committee on Banking Supervision (2010), Basel III: A global regulatory framework for more resilient banks and banking system, Technical report, Bank For International Settlements. December Version.

Bates, J. M. and Granger, C. W. J. (1969), 'The combination of forecasts', *Operations Research Quarterly* **20**, 451–468.

Behringer, J. and van Treeck, T. (2013), Income distribution and current account: A sectoral perspective, IMK Working Paper 125, Macroeconomic Policy Institute.

Belabed, C., Theobald, T. and van Treeck, T. (2013), Income distribution and current account imbalances, IMK Working Paper 126, Macroeconomic Policy Institute.

Bernand, H. and Gerlach, S. (1998), 'Does the term structure predict recessions ? The international evidence', *International Journal of Finance and Economics* **3**, 195–215.

Bernanke, B. (1990), 'On the predictive power of interest rates and interest rate spreads', *New England Economic Review* **Nov./Dec.**, 51–68.

Black, F. and Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–654.

Boldin, M. (1996), 'A check on the robustness of Hamilton's Markov switching model approach to the economic analysis of the business cycle', *Studies in Nonlinear Dynamics and Econometrics* **1**, 35–46.

Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994), ARCH Models, *in* 'Handbook of Econometrics', Vol. 4, R.F. Engle and D.L. McFadden (Eds.), pp. 2959–3038. Amsterdam: North-Holland.

Breunig, R., Najarian, S. and Pagan, A. (2003), 'Specification testing of Markov switching models', *Oxford Bulletin of Economics and Statistics* **65**, 703–725.

Brock, W., Dechert, W., Scheinkman, J. and LeBaron, B. (1996), 'A test for independence based on the correlation dimension', *Econometric Reviews* **15**, 197–235.

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

Burns, A. and Mitchell, W. (1946), *Measuring Business Cycle*, Columbia University Press.

Buss, A., Dumas, B., Uppal, R. and Vilkov, G. (2013), Comparing different regulatory measures to control stock market volatility, Research Grant 2011-2012, Fondation Banque de France.

Campbell, J. Y. and Kyle, A. S. (1993), 'Smart money, noise trading and stock price behavior', *Review of Economic Studies* **60**, 1–34.

Campos, J., Ericsson, N. R. and Hendry, D. (2005), General-to-specific modeling: An overview and selected bibliography, *in* 'General-to-specific Modeling', Edwar-Elgar.

Carrasco, M., Liang, H. and Plogerer, W. (2013), 'Optimal test for markov switching parameters', *Econometrica* **forthcoming**.

Chappe, R. and Semmler, W. (2012), Financial Reform in the U.S.: A Critical Survey of Dodd-Frank and What is Needed for Europe, IMK Studies 28, Macroeconomic Policy Institute.

Chauvet, M. and Potter, S. (2005), 'Forecasting recessions using the yield curve', *Journal of Forecasting* **24**, 77–103.

Chen, S.-H., Lux, T. and Marchesi, M. (2001), 'Testing for nonlinear structure in an artificial financial market', *Journal of Economic Behavior and Organization* **46**, 327–342.

Chen, Z., Iqbal, A. and Lai, H. (2011), 'Forecasting the probability of US recession: A probit and dynamic factor modeling approach', *Canadian Journal of Economics* **44**, 651–672.

Clemen, R. T. (1989), 'Combining forecasts: A review and annotated bibliography', *International Journal of Forecasting* **5**, 559–581.

Colander, D., Föllmer, H., Haas, A., Goldberg, M., Juselius, K., Kirman, A., Lux, T. and Sloth, B. (2009), The Financial Crisis and the Systemic Failure of Academic Economics, Working Paper 1489, Kiel Institute for the World Economy.

Croushore, D. and Stark, T. (2001), 'A real-time data set for macroeconomists', *Journal of Econometrics* **105**, 111 – 130.

Darne, O. and Ferrara, L. (2009), Identification of slowdowns and accelerations for the euro area economy, Documents de Travail 239, Banque de France.

DeGrauwe, P., Dewachter, H. and Embrechts, M. (1993), *Exchange Rate Theory : Chaotic Models of Foreign Exchange Markets*, Blackwell Cambridge.

DeLong, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990), 'Noise trader risk in financial markets', *Journal of Political Economy* **98**, 703–738.

Deutsche Bundesbank (2010*a*), 'Financial Data Deutsche Bundesbank'.
URL: *http://www.bundesbank.de/statistik/statistik_zinsen.php*

Deutsche Bundesbank (2010*b*), 'MFI Balances Sheets Deutsche Bundesbank'.
URL: *http://www.bundesbank.de/statistik/statistik_eszb_tabellen.php*

Deutsche Bundesbank (2010*c*), 'Realtime Data Deutsche Bundesbank'.
URL: *http://www.bundesbank.de/statistik/statistik_realtime.php*

Diebold, F. X. and Rudebusch, G. D. (1991), 'Forecasting output with the composite leading index: A real-time analysis', *Journal of the American Statistical Association* **86**, 603–610.

Drechsel, K. and Scheufele, R. (2012), 'The performance of short-term forecasts of the German economy before and during the 2008/2009 recession', *International Journal of Forecasting* **28**, 428–445.

Dueker, M. (1997), Strenghening the case for the yield curve as a predictor of U.S. recessions, Review, Federal Reserve Bank of St. Louis.

Estrella, A. and Hardouvelis, G. A. (1991), 'The term structure as a predictor of real economic activity', *Journal of Finance* **46**, 555–576.

Estrella, A. and Mishkin, F. S. (1998), 'Predicting U.S. recessions: Financial variables as leading indicators', *Review of Economics and Statistics* **80**, 45–61.

Estrella, A., Rodrigues, A. P. and Schich, S. (2003), 'How stable is the predictive power of the yield curve? Evidence from Germany and the United States', *Review of Economics and Statistics* **85**(3), 629–644.

Fama, E. F. (1970), 'Efficient capital markets: A review of empirical work', *Journal of Finance* **25**, 383–417.

FCI-Commission (2011), The Financial Crisis Inquiry Report, Official Government Edition, Washington D.C.

Friedman, B. M. and Kuttner, K. N. (1992), 'Money, income, prices and interest rates', *American Economic Review* **82**(3), 472–492.

Fritsche, U. and Stephan, S. (2002), 'Leading indicators of German business cycles. an assessment of properties', *Jahrbücher für Nationalökonomie und Statistik* **222**, 289–315.

Gencay, R., Selcuk, F. and Ulugülyagci, A. (2003), 'High volatility, thick tails and extreme value theory in value-at-risk estimation', *Insurance: Mathematics and Economics* **33**, 337 – 356.

Gilchrist, S., Yankov, V. and Zakrajsek, E. (2009), 'Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets', *Journal of Monetary Economics* **56**, 471–493.

Haltmaier, J. (2008), Predicting cycles in economic activity, International Finance Discussion Papers 926, Board of Governors of the Federal Reserve System, Washington D.C.

Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica* **57**(2), 357–384.

Hamilton, J. D. (1990), 'Analysis of time series subject to changes in regime', *Journal of Econometrics* **45**, 39–70.

Hao, L. and Ng, E. C. (2011), 'Predicting Canadian recessions using dynamic probit modeling approaches', *Canadian Journal of Economics* **44**, 1297–1330.

Harding, D. and Pagan, A. (2002), 'Dissecting the cycle: A methodological investigation', *Journal of Monetary Economics* **49**(2), 365–381.

Harvey, C. (1989), 'Forecasts of economic growth from the bond and the stock markets', *Financial Analysts Journal* **Sept./Oct.**, 38–45.

Hellmann, T. F., Murdock, K. C. and Stiglitz, J. E. (2000), 'Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?', *American Economic Review* **90**, 147–165.

Horn, G., Joebges, H., Kamp, L., Krieger, A., Sick, S. and Tober, S. (2009), Gesamtwirtschaftliche Stabilität durch bessere Regulierung - Vorschläge für eine Neuordnung der Finanzmärkte, IMK Report 36, Macroeconomic Policy Institute.

ifo Institut für Wirtschaftsforschung (2011), 'ifo Geschäftsklimaindex', http://www.cesifo-group.de. Accessed on 11 June 2011.

IMK_Konjunkturindikator (2012), 'Macroeconomic Policy Institute'.
URL: *http://www.boeckler.de/imk_38710.htm*

Kahnemann, D. and Tversky, A. (1973), 'On the psychology of prediction', *Psychological Review* **80**, 237–251.

Kauppi, H. and Saikkonen, P. (2008), 'Predicting U.S. recessions with dynamic binary response models', *The Review of Economics and Statistics* **90**(4), 777–791.

Kim, C. (1994), 'Dynamic linear models with Markov - switching', *Journal of Econometrics* **60**, 1–22.

Kirman, A. P. (1992), 'Whom or what does the representative individual represent?', *Journal of Economic Perspectives* **6**, 117–136.

Krolzig, H. M. (1997), *Markov Switching Vector Autoregressions. Statistical Inference and Application to Business Cycle Analysis*, Springer Science+Business Media.

Kumhof, M., Ranciere, R., Lebarz, C., Richter, A. W. and Throckmorton, N. A. (2012), Income inequality and current account imbalances, IMF Working Paper 12/08, International Monetary Fund.

Lahiri, K. and Wang, J. (1994), 'Predicting cyclical turning points with leading index in a Markov switching model', *Journal of Forecasting* **13**(3), 245–263.

Lakonishok, J., Lee, I., Pearson, N. D. and Poteshman, A. M. (2007), 'Option market activity', *Review of Financial Studies* **230**, 814–857.

Lee, C., Liang, C. and Chou, H. (2009), 'Identifying Turning Points of Real Estate Cycles in Taiwan', *Journal of Real Estate Portfolio Management* **15**, 75–85.

Levanon, G. (2010), 'Evaluating and comparing leading and coincident economic indicators', *Business Economics* **45**(1), 16–27.

Lo, A. W. and MacKinlay, A. C. (1988), 'Stock market prices do not follow random walks: Evidence from a simple specification test', *Review of Financial Studies* **1**, 41–66.

Lux, T. (1998), 'The socio-economic dynamics of speculative markets: interacting agents, chaos and the fat tail of return distribution', *Journal of Economic Behavior and Organization* **33**, 143–165.

Lux, T. and Marchesi, M. (2000), 'Volatility clustering in financial markets: A microsimulation of interacting agents', *International Journal of Theoretical and Applied Finance* **3**, 675–702.

Maheu, J. M. and McCurdy, T. H. (2000), 'Identifying bull and bear markets in stock returns', *Journal of Business and Economic Statistics* pp. 100–112.

Malkiel, B. G. (2003), 'The efficient market hyothesis and its critics', *Journal of Economic Perspectives* **17**, 59–82.

Mandelbrot, B. B., Fisher, A. and Calvet, L. (1997), A Multifractal Model of Asset Returns, Cowles Foundation Discussion Paper 1164, Yale University.

Mandelbrot, B. B. and van Ness, J. W. (1968), 'Fractional Brownian Motions, Fractional Noises and Applications', *SIAM Review* **10**, 422–437.

Merton, R. C. (1973), 'Theory of rational option pricing', *The Bell Journal of Economics and Management Science* **4**, 141–183.

Moneta, F. (2005), 'Does the yield spread predict recessions in the Euro area', *International Finance* **8**, 263–301.

National Bureau of Economic Research (2011), Determination of the 2007 Peak in Economic Activity, Technical report, NBER.
URL: *http://www.nber.org/cycles/dec2008.html*

Newey, B. and West, D. (1987), 'A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**, 347–370.

Ng, E. C. (2012), 'Forecasting U.S. recessions with various risk factors and dynamic probit models', *Journal of Macroeconomics* **34**, 112–125.

Nyberg, H. (2010), 'Dynamic probit models and financial variables in recession forecasting', *Journal of Forecasting* **29**(1-2), 215–230.

OECD (2011), 'Realtime Data OECD', http://stats.oecd.org/mei/default.asp?rev=1. Accessed on 11 June 2011.

Pagan, A. (1996), 'The econometrics of financial markets', *Journal of Empirical Finance* **3**, 15–102.

Park, A. and Sabourian, H. (2011), 'Herding and contrarian behavior in financial markets', *Econometrica* **79**, 973 – 1026.

Perlin, M. (2012), MS_Regress - A package for Markov regime switching models in Matlab, SSRN Discussion Paper 1714016, Social Science Research Network.

Pesaran, H. and Timmermann, A. (2005), 'Real time econometrics', *Econometric Theory* **21**, 212–231.

Proaño, C. (2010), Recession forecasting with dynamic probit models under real time conditions, IMK Working Paper 10-2010, Macroeconomic Policy Institute.

Proaño, C. and Theobald, T. (2012), Predicting German recesssions with a composite real-time dynamic probit indicator, NSSR Working Paper 05/12, The New School for Social Research.

Rajan, R. (2010), *Fault Lines: How Hidden Fractures Still Threaten The World Economy*, Princeton University Press.

Repullo, R. (2004), 'Capital requirements, market power, and risk-taking in banking', *Journal of Financial Intermediation* **13**, 156–182.

Ritter, J. R. (2003), 'Behavioral finance', *Pacific-Basin Finance Journal* **11**, 429–437.

Rudebusch, G. D. and Williams, J. C. (2009), 'Forecasting recessions: The puzzle of the enduring power of the yield curve', *Journal of Business and Economic Statistics* **27**(4), 492–503.

Schreiber, S., Theobald, T., Proaño, C., Stephan, S., Rietzler, K. and Detzer, D. (2012), Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeitproblematik - Gutachten im Auftrag des Bundesministeriums der Finanzen (Projekt I A 3 09/10) , IMK Study 27, Macroeconomic Policy Institute.

Shiller, R. J. (2003), 'From efficient market theory to behavioral finance', *Journal of Economic Perspectives* **22**, 83–104.

Shleifer, A. and Summers, L. H. (1990), 'The noise trader approach to finance', *Journal of Economic Perspectives* **4**, 19–33.

Stiglitz, J. E. (2012), *The Price of Inequality: How Today's Divided Society Endangers Our Future*, W.W. Norton and Company New York.

Stock, J. H. and Watson, M. W. (1989), New indexes of coincident and leading economic indicators, *in* 'NBER Macroeconomics Annual', Vol. 4, O. Blanchard and S. Fischer, eds. The MIT Press, Cambridge, MA.

Stock, J. H. and Watson, M. W. (1999), Business cycle fluctuations in U.S. macroeconomic time series, *in* 'Handbook of Macroeconomics', Vol. 1A, J. Taylor and M. Woodford, eds., pp. 3–64. Elsevier.

Theobald, T. (2012), Combining recession probability forecasts from a dynamic probit indicator, IMK Working Paper 89, Macroeconomic Policy Institute.

Theobald, T. (2013), Real-time Markov switching and leading indicators in times of the financial crisis, IMK Working Paper 98, Macroeconomic Policy Institute.

Timmermann, A. (2006), Forecast combinations, *in* 'Handbook of Economic Forecasting', Vol. 1, G. Elliot, C. Granger and A. Timmermann, eds., pp. 135–196. North-Holland, Amsterdam.

Tirole, J. (1985), 'Asset bubbles and overlapping generations', *Econometrica* **53**, 1071–1100.

Wright, J. (2006), The yield curve and predicting recessions, Technical report, Division of Monetary Affairs, Federal Reserve Board.

Ehrenwörtliche Erklärung

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

Düsseldorf, den 20. März 2014