SFB 649 Discussion Paper 2008-064

A note on the model selection risk for ANOVA based adaptive forecasting of the EURIBOR swap term structure

Oliver Blaskowitz*
Helmut Herwatz**

* Humboldt-Universität zu Berlin, Germany
**Christian-Albrechts-Universität zu Kiel, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

A note on the model selection risk for ANOVA based adaptive forecasting of the EURIBOR swap term structure. *

Oliver Blaskowitz
Humboldt–Universität zu Berlin
Institute of Statistics and Econometrics
Phone: +49 – 30 – 2093 – 5705
Email: blaskowitz@wiwi.hu–berlin.de

Helmut Herwartz
Christian–Albrechts–Universität zu Kiel
Institute of Statistics and Econometrics
Phone: +49 – 431 – 880 – 2417
Email: herwartz@stat–econ.uni–kiel.de

Abstract

The paper proposes a data driven adaptive model selection strategy. The selection criterion measures economic ex–ante forecasting content by means of trading implied cash flows. Empirical evidence suggests that the proposed strategy is neither exposed to selection bias nor to the risk of choosing excessively poor models from a parameterized class of candidate specifications.

Keywords: Model selection, Principal components, Factor analysis, Ex–ante forecasting, EURIBOR swap term structure, Trading strategies.
JEL classification: C32, C53, E43, G29.

*This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 'Economic Risk'.

1 Introduction

Quantitative models are formulated to capture regularities governing economic phenomena. As economic relationships are generally complex, models that tempt to describe all aspects of the reality are complicated, too. In order to focus on the essential and to guarantee computational tractability models are formulated which are approximations to real world phenomena. In practice, an analyst frequently encounters a situation in which several candidate models provide reasonable approximations to reality. When selecting a particular model based on a distinct selection criterion, one might be tempted to include as many models as possible into the class of model candidates. Yet, such an approach entails the problem of selection bias. The chance that a model is best in terms of some selection criterion by pure luck increases with the cardinality of the candidate model class (Zucchini 2000). Moreover, with a growing number of models considered the risk of choosing an excessively unfavorable candidate increases.

Within a time series ex–ante forecasting framework we propose a simple method based on an analysis of variance (ANOVA) to select forecasting models out of a parameterized model class. We do not choose a single model for the entire period. Instead, we adaptively select a model in each time instant based on an economic criterion measuring historic ex–ante forecasting performance over a local time window. It is shown that if a forecasting framework is described by a large parameterized model class, the ANOVA based data driven procedure is a suitable tool for model selection which is neither exposed to selection bias nor to the risk of choosing excessively poor models. In the next Section we introduce the class of model candidates, the economic performance criterion and the selection strategy. In Section 3 the empirical results are described using a time series of EURIBOR swap rates. Section 4 concludes.

2 Econometric model and performance measure

The term structure considered is represented by $M = 10$ EURIBOR swap rates with maturities: 3m (3 months), 6m, 1yr (1 year(s)), 2yr, 3yr, 5yr, 7yr, 10yr, 12yr and 15yr. The data is observed over the period from February 22, 1999 to April 25, 2008. To determine rolling

ex–ante forecasts Blaskowitz and Herwartz (2008) employ the model

$$\tilde{y}_t = \Gamma_K F_t + \xi_t, \quad t = T^* - \tau + 1, \ldots, T^*, \tag{2.1}$$

$$\Delta F_t = \nu + \Phi_1 \Delta F_{t-1} + \ldots + \Phi_p \Delta F_{t-p} + \eta_t. \tag{2.2}$$

In (2.1) $\tilde{y}_t = (\tilde{y}_{1t}, \tilde{y}_{2t}, \ldots, \tilde{y}_{Mt})'$ is a vector of swap rates over the $M = 10$ maturities measured in terms of deviations from their local mean, $\tilde{y}_t = y_t - \bar{y}_{T^*}$, $\bar{y}_{T^*} = 1/\tau \sum_{t=T^*-\tau+1}^{T^*} y_t$. F_t is a K–dimensional vector of factors, $F_t = (f_{1t}, ..., f_{Kt})'$, governing the term structure whose changes exhibit vector autoregressive (VAR) dynamics. The M– respectively K–dimensional error terms ξ_t and η_t are treated as white noise processes. The matrix Γ_K is deduced from a principal components analysis (PCA) by decomposing the estimated covariance matrix $\hat{\Sigma}_{T^*}$ of $\tilde{y}_t$ observed over the period $t = T^* - \tau + 1, \ldots, T^*$. More precisely, Γ_K contains the K eigenvectors corresponding to the K largest eigenvalues of $\hat{\Sigma}_{T^*}$.

Conditioning on the information set $\Omega_{T^*,\tau} = \{y_t \mid t = T^* - \tau + 1, \ldots, T^*\}$ h days ahead ex–ante forecasts of swap rates are formalized by means of conditional expectations implied by (2.1) and (2.2). This gives

$$\hat{y}_{T^*+h} = \Gamma_K \hat{F}_{T^*+h} + \bar{y}_{T^*},$$

where factor implied forecasts are readjusted for the local in–sample mean. To compute factor forecasts $\hat{F}_{T^*+h}$ we estimate a VAR(p) model based on extracted factors $\hat{F}_t = \Gamma_K' \tilde{y}_t$ and iterate forecasts h days ahead.

Model performance is measured by trading implied cash flows. In particular, we consider three strategies based on daily ex–ante forecasts for the 2yr, 5yr and 10yr swap rate. For h days ahead predicted upward (downward) movements in swap rates payer (receiver) swap agreements with a notional of 100 Euros are entered. Cash flows for the swap positions are computed by means of the comparison swap valuation technique (Miron and Swannell 1991) using the realized 2yr, 5yr, 10yr swap rates h days ahead. The bank account starts at zero and contains cumulated cash flows inferred from daily forecasts up to time T^*. Total cash flows are given by cumulated cash flows at the end of the trading period.

3 Results

When specifying the parameters τ, K and p the analyst faces a model selection problem. Blaskowitz and Herwartz (2008) consider a plentitude of 100 contending PCA/VAR models

implied by alternative parameter selections $\tau \in \{42, 63, 126, 189, 252\}$, $K \in \{1, 2, 3, 4, 5\}$ and $p \in \{0, 1, 2, 3\}$. They propose an adaptive model selection procedure based on an ANOVA of trading implied cash flows. The latter procedure is implemented by regressing total cash flows generated over a second local window of 42 days on dummy variables related to the model characteristics (τ, K, p). In this section we compare the economic forecast performance of $h = 1, 5, 10$ days ahead ex–ante forecasts of the adaptive approach and the set of unconditional models. To ensure that the forecasts of all models are compared over the same period of swap rates the rolling window analysis starts on April 19, 2000. Hence 2083 ex–ante forecasts for the 100 unconditional models and the adaptive strategy are available.

From the top panel of Table 1 it can be seen that all adaptive strategies generate positive total cash flows over the 2083 trading days. In particular, total cash flows increase with the forecast horizon as swap rate movements are likely to be larger over multiple days. The performance of the adaptive model selection strategies compares favorably to the performance of the unconditional modeling approach. In four cases out of nine (2yr and 5yr for $h = 5$ and $h = 10$ resp.) the adaptive strategy outperforms at least 90 out of 100 unconditional models. Moreover, for all forecast exercises the adaptive approach is better than at least 70 unconditional specifications. Furthermore, the total cash flow of the adaptive strategy exceeds the average cash flow over all unconditional PCA/VAR models for all trades considered. The class of unconditional PCA/VAR models considered contains specifications having considerable downside risk. The model specification with the minimum total cash flow (MinCF) always generates negative cash flows. Yet, the adaptive selection procedure seems to filter out unfavorable models as indicated by the corresponding total cash flows.

The above ranking of total cash flows is only a snapshot characterizing the last trading day. Figure 1 shows the ranking of the adaptive strategy with respect to unconditional models in terms of cumulated cash flows over time. The data driven ANOVA selection procedure performs strongly in 6 out of 9 (5yr and 10yr for $h = 1$, 2yr and 5yr for $h = 5$ and $h = 10$ resp.) forecast exercises. In these cases the adaptive strategy cumulates higher cash flows than at least 50 unconditional models. Particularly convincing is the performance of the adaptive procedure for the 2yr and 5yr trades for horizons $h = 5, 10$. It performs better than at least 80 unconditional models. For the three remaining forecast exercises (2yr for $h = 1$, 10yr for $h = 5, 10$) it can be seen that in terms of cumulated cash flows the ranking of the

ANOVA selection procedure is more or less fluctuating around the median of unconditional models. The last line of Table 1 shows that the adaptive strategy's time average of the ranking is larger than the time average of the ranking of 50, 69 and 65 unconditional models resp.

Table 1 also provides an insight into the evolution of the bank account for the adaptive and unconditional strategies over the 2083 trading days. The minimum bank account of the adaptive model selection procedure is negative for all forecast exercises considered as can be seen from the first panel of Table 1. However the adaptive approach is always better than at least 50 of the 100 unconditional models as shown in the third panel of Table 1 (RankMinBank).

	2yr			5yr			10yr		
	$h=1$	$h=5$	$h=10$	$h=1$	$h=5$	$h=10$	$h=1$	$h=5$	$h=10$
CF	12.24	38.84	100.18	33.49	71.00	164.41	36.42	78.94	118.88
MinBank	-1.29	-3.26	-7.74	-1.13	-8.12	-20.42	-1.13	-14.15	-29.30
MinCF	-1.00	-5.98	-24.80	-11.58	-47.13	-53.56	-12.07	-90.05	-106.08
AvCF	9.86	18.38	31.97	18.33	32.42	58.84	26.08	51.01	80.49
RankCF	70	90	100	88	92	98	80	79	78
MaxCF	22.31	46.91	95.33	45.05	96.51	180.49	48.23	138.51	244.49
MinMinBank	-5.35	-21.72	-46.44	-12.41	-55.28	-71.45	-13.69	-105.53	-144.88
RankMinBank	51	77	74	48	68	74	73	54	53
MaxMinBank	0.03	-0.85	-3.16	0.20	-1.36	-10.46	0.15	-0.29	-6.37
AvRankBank	50	87	94	83	91	100	88	69	65

Table 1. Results for the adaptive strategy and some summary measures for the set of unconditional models for each forecast exercise over the period April 19, 2000 to April 11, 2008. CF and MinBank denote the total cash flow and the minimum bank account of the adaptive strategy. MinCF, AvCF and MaxCF give the minimum, average and maximum total cash flow, respectively, over all the unconditional PCA/VAR models. RankCF is the rank of the adaptive strategy within the set of unconditional models when ordered according to total cash flows. MinMinBank and MaxMinBank denote the minimum resp. maximum value of the minimum bank account over the 2083 trading days over all unconditional models. RankMinBank resp. AvRankBank are the ranks of the adaptive strategy within the set of unconditional models when ordered according to the minimum bank account value resp. when ordered according to the time average of ranks (see also Figure 1).

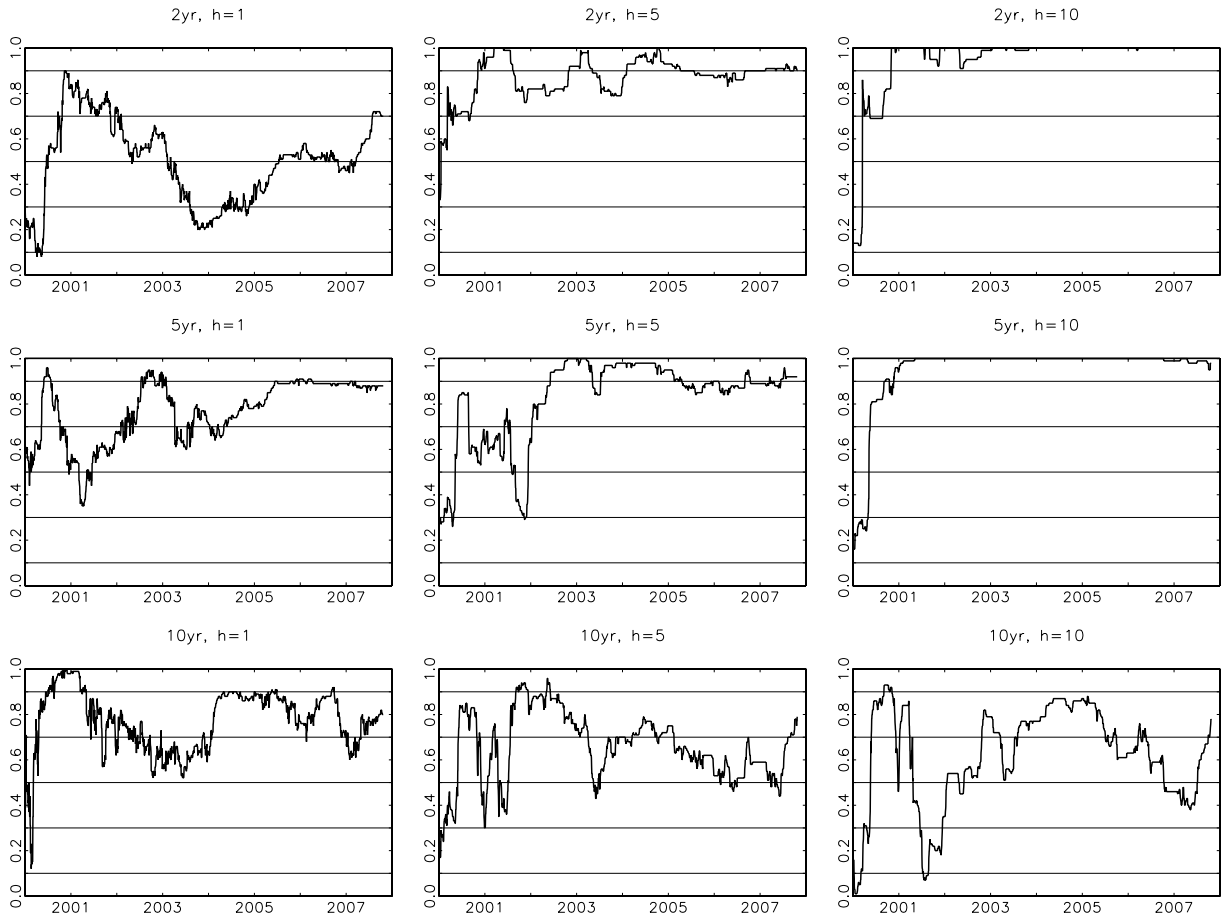


Figure 1. Rankings of the adaptive strategy with respect to the set of 100 unconditional models over the period April 19, 2000 to April 11, 2008. The value of the bank account at time T^* is used as the ranking criterion. Horizontal lines represent the 10th, 30th, ..., 90th quantile of unconditional models in time T^* according to the bank account value.

4 Conclusion

In an empirical application to forecast the EURIBOR swap term structure we show that ANOVA based adaptive model selection performs strongly when compared to a set of 100 candidate models in economic terms. As a conclusion data–driven adaptive model selection based on local ex–ante economic forecasting performance merits consideration to avoid model selection bias.

References

Blaskowitz, O. and H. Herwartz, 'PCA based ex–ante forecasting of swap term structures', *International Journal of Theoretical and Applied Finance*, forthcoming in 2008.

Miron P. and P. Swannell, *Pricing and hedging swaps*, Euromoney Publications PLC, London (1991).

Zucchini, W., 'An introduction to model selection', *Journal of Mathematical Psychology*, **44**, 41–61 (2000).

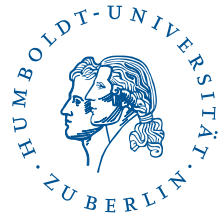
SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.
007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zinych, January 2008.
017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".



022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.
027 "The Stochastic Fluctuation of the Quantile Regression Curve" by Wolfgang Härdle and Song Song, March 2008.
028 "Are stewardship and valuation usefulness compatible or alternative objectives of financial accounting?" by Joachim Gassen, March 2008.
029 "Genetic Codes of Mergers, Post Merger Technology Evolution and Why Mergers Fail" by Alexander Cuntz, April 2008.
030 "Using R, LaTeX and Wiki for an Arabic e-learning platform" by Taleb Ahmad, Wolfgang Härdle, Sigbert Klinke and Shafeeqah Al Awadhi, April 2008.
031 "Beyond the business cycle – factors driving aggregate mortality rates" by Katja Hanewald, April 2008.
032 "Against All Odds? National Sentiment and Wagering on European Football" by Sebastian Braun and Michael Kvasnicka, April 2008.
033 "Are CEOs in Family Firms Paid Like Bureaucrats? Evidence from Bayesian and Frequentist Analyses" by Jörn Hendrich Block, April 2008.
034 "JBendge: An Object-Oriented System for Solving, Estimating and Selecting Nonlinear Dynamic Models" by Viktor Winschel and Markus Krätzig, April 2008.
035 "Stock Picking via Nonsymmetrically Pruned Binary Decision Trees" by Anton Andriyashin, May 2008.
036 "Expected Inflation, Expected Stock Returns, and Money Illusion: What can we learn from Survey Expectations?" by Maik Schmeling and Andreas Schrimpf, May 2008.
037 "The Impact of Individual Investment Behavior for Retirement Welfare: Evidence from the United States and Germany" by Thomas Post, Helmut Gründl, Joan T. Schmit and Anja Zimmer, May 2008.
038 "Dynamic Semiparametric Factor Models in Risk Neutral Density Estimation" by Enzo Giacomini, Wolfgang Härdle and Volker Krätschmer, May 2008.
039 "Can Education Save Europe From High Unemployment?" by Nicole Walter and Runli Xie, June 2008.
040 "Solow Residuals without Capital Stocks" by Michael C. Burda and Battista Severgnini, August 2008.
041 "Unionization, Stochastic Dominance, and Compression of the Wage Distribution: Evidence from Germany" by Michael C. Burda, Bernd Fitzenberger, Alexander Lembcke and Thorsten Vogel, March 2008
042 "Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungs-methodik" by Dirk Temme and Lutz Hildebrandt, June 2008.
043 "Modeling Dependencies in Finance using Copulae" by Wolfgang Härdle, Ostap Okhrin and Yarema Okhrin, June 2008.
044 "Numerics of Implied Binomial Trees" by Wolfgang Härdle and Alena Mysickova, June 2008.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

045 "Measuring and Modeling Risk Using High-Frequency Data" by Wolfgang Härdle, Nikolaus Hautsch and Uta Pigorsch, June 2008.
046 "Links between sustainability-related innovation and sustainability management" by Marcus Wagner, June 2008.
047 "Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models" by Nikolaus Hautsch and Vahidin Jeleskovic, July 2008.
048 "Macro Wine in Financial Skins: The Oil-FX Interdependence" by Enzo Weber, July 2008.
049 "Simultaneous Stochastic Volatility Transmission Across American Equity Markets" by Enzo Weber, July 2008.
050 "A semiparametric factor model for electricity forward curve dynamics" by Szymon Borak and Rafał Weron, July 2008.
051 "Recurrent Support Vector Regreson for a Nonlinear ARMA Model with Applications to Forecasting Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang K. Härdle, July 2008.
052 "Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality" by Wolfgang Reichmuth and Samad Sarferaz, July 2008.
053 "Yield Curve Factors, Term Structure Volatility, and Bond Risk Premia" by Nikolaus Hautsch and Yangguoyi Ou, July 2008.
054 "The Natural Rate Hypothesis and Real Determinacy" by Alexander Meyer-Gohde, July 2008.
055 "Technology sourcing by large incumbents through acquisition of small firms" by Marcus Wagner, July 2008.
056 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycle" by Fang Yao, August 2008.
057 "Measuring changes in preferences and perception due to the entry of a new brand with choice data" by Lutz Hildebrandt and Lea Kalweit, August 2008.
058 "Statistics E-learning Platforms: Evaluation Case Studies" by Taleb Ahmad and Wolfgang Härdle, August 2008.
059 "The Influence of the Business Cycle on Mortality" by Wolfgang H. Reichmuth and Samad Sarferaz, September 2008.
060 "Matching Theory and Data: Bayesian Vector Autoregression and Dynamic Stochastic General Equilibrium Models" by Alexander Kriwoluzky, September 2008.
061 "Eine Analyse der Dimensionen des Fortune-Reputationsindex" by Lutz Hildebrandt, Henning Kreis and Joachim Schwalbach, September 2008.
062 "Nonlinear Modeling of Target Leverage with Latent Determinant Variables – New Evidence on the Trade-off Theory" by Ralf Sabiwalsky, September 2008.
063 "Discrete-Time Stochastic Volatility Models and MCMC-Based Statistical Inference" by Nikolaus Hautsch and Yangguoyi Ou, September 2008.
064 "A note on the model selection risk for ANOVA based adaptive forecasting of the EURIBOR swap term structure" by Oliver Blaskowitz and Helmut Herwartz, October 2008.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".