

Baruník, Jozef; Kukačka, Jiří

**Working Paper**

## Simulated ML Estimation of Financial Agent-Based Models

IES Working Paper, No. 07/2016

**Provided in Cooperation with:**

Charles University, Institute of Economic Studies (IES)

*Suggested Citation:* Baruník, Jozef; Kukačka, Jiří (2016) : Simulated ML Estimation of Financial Agent-Based Models, IES Working Paper, No. 07/2016, Charles University in Prague, Institute of Economic Studies (IES), Prague

This Version is available at:

<https://hdl.handle.net/10419/174174>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Institute of Economic Studies, Faculty of Social Sciences  
Charles University in Prague

# Simulated ML Estimation of Financial Agent-Based Models

Jiri Kukacka  
Jozef Barunik

IES Working Paper: 07/2016



Institute of Economic Studies,  
Faculty of Social Sciences,  
Charles University in Prague

[UK FSV – IES]

Opletalova 26  
CZ-110 00, Prague  
E-mail : [ies@fsv.cuni.cz](mailto:ies@fsv.cuni.cz)  
<http://ies.fsv.cuni.cz>

Institut ekonomických studií  
Fakulta sociálních věd  
Univerzita Karlova v Praze

Opletalova 26  
110 00 Praha 1

E-mail : [ies@fsv.cuni.cz](mailto:ies@fsv.cuni.cz)  
<http://ies.fsv.cuni.cz>

**Disclaimer:** The IES Working Papers is an online paper series for works by the faculty and students of the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Czech Republic. The papers are peer reviewed, but they are *not* edited or formatted by the editors. The views expressed in documents served by this site do not reflect the views of the IES or any other Charles University Department. They are the sole property of the respective authors. Additional info at: [ies@fsv.cuni.cz](mailto:ies@fsv.cuni.cz)

**Copyright Notice:** Although all documents published by the IES are provided without charge, they are licensed for personal, academic or educational use. All rights are reserved by the authors.

**Citations:** All references to documents served by this site must be appropriately cited.

**Bibliographic information:**

Kukacka J., Barunik, J. (2016). “Simulated ML Estimation of Financial Agent-Based Models” IES Working Paper 7/2016. IES FSV. Charles University.

This paper can be downloaded at: <http://ies.fsv.cuni.cz>

# Simulated ML Estimation of Financial Agent-Based Models

Jiri Kukacka<sup>a,b,c</sup>

Jozef Barunik<sup>a,b</sup>

<sup>a</sup>Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Smetanovo nabrezi 6, 111 01 Prague 1, Czech Republic

<sup>b</sup>Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodarenskou vezi 4, 182 00 Prague 8, Czech Republic

<sup>c</sup>Email (corresponding author): [jiri.kukacka@fsv.cuni.cz](mailto:jiri.kukacka@fsv.cuni.cz)

March 2016

## Abstract:

This paper proposes computational framework for empirical estimation of Financial Agent-Based Models (FABMs) that does not rely upon restrictive theoretical assumptions. We customise a recent methodology of the Non-Parametric Simulated Maximum Likelihood Estimator (NPSMLE) based on kernel methods by Kristensen and Shin (2012) and elaborate its capability for FABMs estimation purposes. To start with, we apply the methodology to the popular and widely analysed model of Brock and Hommes (1998). We extensively test finite sample properties of the estimator via Monte Carlo simulations and show that important theoretical features of the estimator, the consistency and asymptotic efficiency, also hold in small samples for the model. We also verify smoothness of the simulated log-likelihood function and identification of parameters. Main empirical results of our analysis are the statistical insignificance of the switching coefficient but markedly significant belief parameters defining heterogeneous trading regimes with an absolute superiority of trend-following over contrarian strategies and a slight proportional dominance of fundamentalists over trend following chartists.

**Keywords:** Heterogeneous Agent Model, Heterogeneous Expectations, Behavioural Finance, Intensity of Choice, Switching, Non-Parametric Simulated Maximum Likelihood Estimator.

**JEL:** C14, C51, C63, D84, G02, G12.

**Acknowledgements:** This paper benefited much from thoughtful comments by E. Gerba, L. Vacha, and R.C.J. Zwinkels. We also appreciate efforts of organisers of WEHIA 2014 and 2015, CFE 2014, Econophysics Colloquium 2014, First Bordeaux-Milano Joint Workshop on A-B Macroeconomics 2015, and CEF 2015 and we are



grateful to discussants for many constructive remarks. Especially we wish to thank to C. Hommes, B. LeBaron, and S. Barde for their discussion on the very recent version of the paper during CEF 2015 in Taipei. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. FP7-SSH- 612955 (FinMaP). Support from the Czech Science Foundation under the P402/12/G097 DYME - “Dynamic Models in Economics” project is gratefully acknowledged. J. Kukacka gratefully acknowledges financial support from the Grant Agency of Charles University under the 192215 project.

## 1. Introduction

Financial markets are one of the fundamental motivative forces of the economic development but the Global Financial Crisis pointed again at the deficiency of knowledge of how this important segment of the global economy works. After the failure of traditional financial models in the Global Financial Crisis of 2007–2008, the Agent-Based (AB) approaches in Finance denoted as Financial Agent-Based Models (FABMs) have attracted attention both of academicians as well as practitioners and hence gradually replace traditional financial models in the recent financial literature. This advancement emphasises that although the serious macroeconomic consequences of market fluctuations are worldwide, the essence of problems remains at the level of individual market agents with their heterogeneous expectations. The FABMs reflect this well documented and systematic human departure from the representative agent’s full rationality towards reasonably realistic bounded, limited rationality (Simon, 1957). An essential achievement of the FABM methodology is the ability to replicate so called stylised facts of financial data<sup>1</sup> and account for emergence of asset market bubbles followed by sudden crashes. Neither observed empirical regularities, nor explosive bubbles (Evans, 1991) can be reasonably explained by traditional financial models. Recently, number of projects propose a courageous attempt to complement or even alternate current mainstream policy making approaches through the use of Agent-Based Models (ABMs), typically at the level of central banks. For this to happen, it is, however, essential to estimate these models on the empirical data in order to use them for forecasting.

Traditional models in Economics and Finance are based on the hypothesis of Rational Expectations (RE) (Muth, 1961; Lucas, 1972) and approximation of market population by a representative agent. Under RE, agents form expectation using all available information, however, they may be individually incorrect. Nonetheless, agents must not be systematically biased, i.e. the forecasting errors agents make must be random. The representative agent, which notion dates back to Edgeworth (1881), thus behaves in a perfectly rational, i.e. model consistent manner according to solution of a maximisation problem under full information (involving also information about behaviour of all other agents) and no computational constraints. Especially in Finance, mostly simple linear, stable equilibrium models driven by exogenous random news about fundamentals have been developed under this paradigm. A ‘textbook example’ is the Capital Asset Pricing Model (CAPM, e.g. Sharpe, 1964). RE is a necessary condition for the striking Efficient Market Hypothesis (EMH) (Fama, 1970), dominating the field in the past, according to which asset prices reflect all relevant information about economic fundamentals available to economic agents. As a consequence, securities prices follow Random Walk (RW). Irrational traders thus in average receive lower profits than rational agents and in the process of the ‘evolutionary market pressure’ are driven out of the market, a statement called the ‘Friedman Hypothesis’.

This paper focuses on the field of Agent-Based Computational Finance (ACF) that has experienced an extensive development during the last three decades. The departure of FABMs from the RE paradigm has proceeded from the 1980s ensued by first macroeconomic ABMs from the 1990s. Recently, many Macro ABMs have been developed sharing similar modelling concepts with FABMs but also following the Dynamic Stochastic General Equilibrium (DSGE) literature as many challenges within these two fields overlap. A rapid development of Macro ABMs was substantially accelerated by events in 2008, known as the Global Financial Crisis of 2007–2008, followed by the

---

<sup>1</sup>A term coined by Kaldor (1961, pg. 178) as view of the facts concentrated “on broad tendencies, ignoring individual detail”, for comprehensive surveys consult Cont (2001, 2007).

period of so called ‘Great Recession’. [Fagiolo and Roventini \(2012, pg. 67, 69\)](#) comment that “the Great Recession seems to be a natural experiment for macroeconomics showing the inadequacy of the predominant theoretical framework—the New Neoclassical Synthesis—grounded on the DSGE model” and draw attention to the fact that “an increasing number of leading economists claim that the current economic crisis is a crisis for economic theory”. [Canova et al. \(2014, pg. 1029\)](#) argue that “linear Gaussian specifications [of DSGE models]<sup>2</sup> are inadequate to describe the 2008–2009 Great Recession, the subsequent episode of zero nominal interest rates and the events during the subsequent sovereign debt crisis in Europe”.

AB approaches in Economics thus departure from models with perfectly rational representative agent<sup>3</sup> and model-consistent homogeneous expectations in reaction to unrealistic assumptions of the RE paradigm, equilibrium conditions, lack of microeconomics foundation when applying the RE Hypothesis in macroeconomic research,<sup>4</sup> and the inability of asset pricing models derived from the EMH to replicate empirically observed stylised facts and explain speculative bubbles. This modern approach builds on direct interactions of boundedly rational economic agents ([Simon, 1955, 1957; Sargent, 1993](#)) with limited cognitive and information processing capacities, disposing insufficient computational power, and incomplete information. Nonetheless, agents do not act irrationally, but follow simple behavioural heuristics, that may be the most ‘rational’ choice given objective constraints they face and costs of gathering information. Agents’ actions are not solutions of maximisation problems, but are selected according to adaptive updating rules and their relative profitability. According to [Branch \(2004, pg. 592\)](#), they in fact “behave as if they were econometricians”. Agents are assumed to behave according to psychological and sociological evidence to better reflect the real world phenomena, i.e. ABMs often embrace findings from market psychology (e.g. [Kahneman and Tversky, 1974, 1979](#)) and herding behaviour ([Keynes, 1936](#)). Dynamics of these economic systems is not generated via exogenous shock mechanisms but prices are driven endogenously based on boundedly rational expectations of agents resulting in direct interactions. Any equilibrium condition is not required, that means, markets may be found even continually out of equilibrium without violating model assumptions. Another important theoretical viewpoint in favour of ABMs is revealed e.g. by [Browning et al. \(1999\)](#) who remark that representative preferences mostly cannot govern model behaviour asymptotically, or by [Fagiolo et al. \(2008\)](#) who rightly point out that RE prevent models to address distributional issues in situations when many macroeconomic time series distributions can be well-approximated by fat tail densities.

Although the empirical estimation is an important part of the modelling cycle and seems crucial for model validation, one cannot find many attempts on empirical estimation of FABMs. Moreover, looking ten years back in the financial literature, we neither observe any general consensus on the estimation methodology, nor conclusive results. [Fagiolo et al. \(2007, pg. 202\)](#) even emphasise “no consensus at all about how (and if) AB models should be empirically validated”. Generally, there are two essential difficulties, or rather challenges, in estimating the FABMs. First, a highly nonlinear and complex nature of these systems prohibits researchers of using classical estimation methods as the objective function often has no analytical expression. Second, a possible overparametrisation, high number of degrees of freedom, and optional model settings together with

---

<sup>2</sup>A note added by the authors.

<sup>3</sup>An important early criticism of the representative agent paradigm is provided by [Kirman \(1991\)](#).

<sup>4</sup>So called ‘Aggregation Problem’ refers to a theoretical fact that the assumption of rationality at the individual level does not imply aggregate rationality (e.g. [Janssen, 1993](#)). [Fagiolo and Roventini \(2012\)](#) correctly point out that “RE is a property of the economic system as a whole, individual rationality is not a sufficient condition for letting the system converge to the RE fixed-point equilibrium ([Howitt, 2012](#))”.

the stochastic dynamics further escalate the complexity of the problem. The emerging properties of these models cannot be analytically deduced, a Method of Moments (MM), “while fine in theory, might be too computationally costly to undertake” (LeBaron and Tesfatsion, 2008, pg. 249), and thus a considerable simulation capacity for the numerical analysis is required.

This paper makes a step forward and proposes more general computational framework for empirical validation of full-fledged FABMs. We base the estimation methodology on a recently developed Non-Parametric Simulated Maximum Likelihood Estimator (NPSMLE) by Kristensen and Shin (2012). For many FABMs one cannot analytically derive the likelihood function to estimate the model parameters via Maximum Likelihood Estimator (MLE). However, the observations from the model can be numerically simulated and utilised for the kernel estimation of the conditional density of the data-generating process. Thus, the likelihood function can be replaced by the simulated likelihood. Non-Parametric Simulated Maximum Likelihood Estimator (NPSMLE) is an estimation framework that functions under very general conditions met by many FABMs. Hence its theoretical properties hold and it can be transferred to the FABM literature. Indeed, recently Grazzini and Richiardi (2015, pg. 151) suggest to employ the NPSMLE methodology on ABMs in general. We extensively test capability of the method for the FABMs estimation purposes via a complex Monte Carlo analysis. To start with, we apply the methodology to the popular and widely analysed model of Brock and Hommes (1998) for which we customise the general framework of Kristensen and Shin (2012). The key feature of the model is an evolutionary switching of agents between simple trading strategies based on past realised profits—so called Adaptive Belief System (ABS)—governed by the switching parameter of the intensity of choice  $\beta$ . This parameter is responsible for high nonlinearity of the system and possibly chaotic price motion. We presuppose that if the NPSMLE method succeeds in estimation of this generally challenging FABM framework and the switching parameter  $\beta$ , it is likely to appear more general and useful for other ABMs in the future.

The paper is organised as follows. After the Introduction, in Section 2 we provide a literature survey on FABMs estimation methods. Next, Section 3 introduces theoretical background of the NPSMLE method by Kristensen and Shin (2012) and describes the Brock and Hommes (1998) Heterogeneous Agent Model (HAM) framework. Findings of the Monte Carlo simulation study of NPSMLE application to the HAM are reported in Section 4 and Section 5 presents empirical estimation results. Finally, Section 6 draws overall conclusions.

## 2. Literature review: methods and results

Over last decades, a large number of various HAMs have been developed and analysed. However, although the empirical estimation is an important validation part of the modelling cycle, one cannot find many examples on empirical estimation of HAMs using empirical data—typical HAM studies mostly employ simulation techniques to confirm ability to replicate stylised facts of financial data. Additionally, only several of those attempts provide a rigorous comparison of forecasting performance or in terms of fitting empirical market data with ‘mainstream’ approaches such as ARIMA, GARCH ‘family’ or other ‘competing’ econometrics models. In existing empirical papers, estimation methods are often chosen ad hoc or the models are ex ante designed or substantially simplified in a way that a particular estimation method can be used. For the reason, as de Jong et al. (2010, pg. 1653) point out: “although the heterogeneity of agents approach is intellectually satisfying, the heterogeneity model has hardly been estimated with empirical financial data because of the non-linear nature of the model that mainly arises from the existence of the mechanism that

governs the switching between beliefs”. Furthermore, [Westerhoff and Reitz \(2005, pg. 642\)](#) highlight the fact that “one has to sacrifice certain real-life market details. If the setup is too complicated, econometric analysis is precluded”. In any case, since the complexity of HAMs often does not allow for analytical solutions, the empirical validation of agent-based systems together with simulation analyses remain the crucial tools of HAMs verification.

### *2.1. The use of econometric techniques*

Utilisation of econometrics to empirically validate or estimate HAMs dates more than one decade back in the financial literature history. Within this stage of development of the Heterogeneous Agent (HA) modelling, the central concern embraces the determination of appropriate values of model parameters and assessment of their statistical significance. However, as summarised in [Table 1](#) and [Table 2](#), looking ten years back in literature, we neither observe any general consensus on the estimation methodology, nor conclusive results. [Fagiolo et al. \(2007, pg. 199, 202\)](#) assert that “a strongly heterogeneous set of approaches to empirical validation is to be found in the AB literature”. Given different origins as well as various modelling concepts, the estimation methodology also varies. As depicted in the third column of [Tables 1](#) and [2](#), the three estimation methods—the Nonlinear Least Squares (NLS), Quasi Maximum Likelihood (QML), and the Method of Simulated Moments (MSM)—prevail among others. When moving to the fourth column of [Tables 1](#) and [2](#) we can see how the choice of estimated parameters is affected by various model designs. Nonetheless, we can observe a general tendency to estimate mainly parameters related to ‘behavioural rules’ of agents: belief coefficients defining individual trading strategies and the intensity of choice or its corresponding concepts in different types of models (mutation, herding tendency, and switching thresholds). All these parameters are apparently meaningful from the economic interpretation point of view.

Various direct and indirect estimation methods have already been employed. However, for the use of direct methods, instead of the usual Ordinary Least Squares (OLS) or Maximum Likelihood (ML) methods, the NLS and QML methods are applied in most of the cases. In these applications, crucial HAM structural features, e.g. the evolutionary switching between trading strategies—one of the key concepts of the HA modelling, are sometimes restrained or even sacrificed to obtain simplified approach which can be estimated using suggested methods. However, for many HAMs the aggregation equation, which would contain all parameters of interest, cannot be derived analytically and therefore the application of direct estimation techniques is not feasible. Indirect estimation methods thus overcome this problematic issue by simulating artificial data from the model through which the aggregation concepts such as moments for the MSM are derived. These simulation-based econometric methods “are very applicable and may dramatically open the empirical accessibility of agent-based models in the future” as suggested by [Chen et al. \(2012, pg. 204\)](#). Simulation-based econometric methods already used for the HAMs estimation include the MSM, the Efficient Method of Moments (EMM), or generally the Simulated Minimum Distance (SMD). All these methods are based on minimising the (weighted) distance between two sets of simulated and observed moments. So far, however, the use of simulation-based econometric methods for validation of HAMs is relatively rare.

### *2.2. Performance of ML and Quasi ML*

Applications of methods based on ML principle share a relatively similar problem: the objective function is often very flat in direction of some parameters—typically the switching parameter of the intensity of choice. Problematic identification of given parameters is then reflected in large

Table 1: Estimation methods of FABMs I.

Models	Origin	Methods	Parameters estimated	#	Data	Type	Fit	IOC
Alfarano et al. (2005)	IAH	ML	Herding tendency	2	d:5034–9761 o.	s,fx,g	-	-
Alfarano et al. (2006)	IAH	ML	Herding tendency	2	d:5495,6523 o.	s,fx	-	-
Alfarano et al. (2007)	IAH	ML	Herding tendency	2	d:1975–2001	s	-	-
Amilon (2008)	ABS	EMM/ML	Intensity of choice <sup>a</sup>	15	d:1980–2000	s	$p\text{-}v=0\%$	1.99(i),1.91(s)
Boswijk et al. (2007)	ABS	NLS	Belief coefficients/Intensity of choice	3	a:132 o.	s	$R^2=.82$	10.29(i),7.54(i)
de Jong et al. (2009b)	ABS	NLS	Belief coefficients/Intensity of choice	5	w:102 o.	fx	$adj R^2=.14$	1.52(i)
de Jong et al. (2010)	ABS	Quasi ML	Belief coefficients/Intensity of choice	7	m:238 o.	fx	-	.0007(i)–6.29(s)
Diks and Weide (2005)	ABS	ML	(G)ARCH relations/Sign of MA(1) c.	3	d:3914 o.	fx	-	-
Ecemis et al. (2005)	AA	IEC	Market fractions/Behavioural rules	3	-	s	-	-
Gilli and Winker (2003)	ANT	MSM	Mutation/Conviction rate	3	d:1991–2000	fx	NA	-
Manzan and Westerhoff (2007)	ABS	OLS	Reaction coeffs./Switching threshold	4	m:1/74–12/98	fx	NA	-
Reitz and Westerhoff (2007)	ABS	Quasi ML	Behavioural rules/Intensity of choice	6	m:365 o.	c	-	.17(s)–.47(s)
Westerhoff and Reitz (2003)	ABS	Quasi ML	Behavioural rules/Intensity of choice	7	d:4431 o.	fx	-	.02(s)–.17(s)
Winker and Gilli (2001)	ANT	MSM	Mutation/Conviction rate	2	d:1991–2000	fx	NA	-
Winker et al. (2007)	ANT	MSM	Mutation/Conviction rate	3	d:1991–2000	fx	$p\text{-}v<1\%$ <sup>b</sup>	-

*Note:* The Table is adopted from [Chen et al. \(2012, pg. 203\)](#) and amended by the authors. Authors are alphabetised. The full meaning of the acronyms under ‘Origin’: AA stands for Autonomous Agents, ABS for Adaptive Belief System, ANT for the Ant type of system, and IAH for Interactive Agent Hypothesis. The full meaning of the acronyms under ‘Methods’: ML stands for Maximum Likelihood, EMM for Efficient Method of Moments, NLS for Nonlinear Least Squares, OLS for Ordinary Least Squares, IEC for Interactive Evolutionary Computation, and MSM (SMM) for Method of Simulated Moments. ‘#’ displays total number of estimated parameters; ‘Data’ describes data frequency: ‘d/w/m/q/a’ for daily/weekly/monthly/quarterly/annual, and number of observations (when a specific figure is not provided, we report starting and final years); ‘Type’ shows the type of data: ‘s/fx/c/g/re’ for stock markets/FX/commodity markets/gold/real estate; ‘Fit’ reports the statistical fit of the estimation ( $R^2$ , its alternatives, p-value of the J-test of overidentifying restrictions to accept the model as a possible data generating process); and ‘|IOC|’ displays the absolute estimated value of the ‘intensity of choice’—the switching parameter from the multinomial logit model, see [Equation 16](#) (where relevant), furthermore ‘s’/‘i’ denotes its statistical significance/insignificance at 5% level. Figures are rounded to 2 decimal digits.

<sup>a</sup>[Chen et al. \(2012\)](#) do not report other important parameters estimated: belief coefficients, intensities of exogenous noises, risk aversion, information costs for fundamentalists, forgetting factors, and memory in the fitness measure.

<sup>b</sup>While p-val for GARCH(1,1) model > 5%.

Table 2: Estimation methods of FABMs II. a)

Models	Origin	Methods	Parameters estimated
Barunik and Vosvrda (2009)	Cusp	ML	Asymmetry and bifurcation factors/Location and scale coefficient
Barunik and Kukacka (2015)	Cusp	RV/ML	Asymmetry and bifurcation factors/Polynomial data approximation
Bolt et al. (2011)	ABS	NLS(?)	Expectations' bias/Discount factor/Belief coefficients/Intensity of choice
Bolt et al. (2014)	ABS	NLS	Belief coefficients/A-synchronous updating ratio/Intensity of choice
Cornea et al. (2013)	ABS	VAR/NLS	Fundamentalists' belief coefficient/Intensity of choice
Chen and Lux (2015)	IAH	MSM	Standard deviation of innovations/Herding tendency
Chiarella et al. (2014)	ABS	Quasi ML	Belief & market maker coefficients/Memory decay rate/Intensity of choice
Chiarella et al. (2015)	ABS	Quasi ML	Belief coefficients/Variance risk premium/Intensity of choice
de Jong et al. (2009a)	ABS	ML	Belief coefficients/Intensity of choice
ter Ellen and Zwinkels (2010)	ABS	Quasi ML	Belief coefficients/Intensity of choice
ter Ellen et al. (2013)	ABS	OLS/NLS	Behavioural rules/Intensity of choice
Franke (2009)	ABS	MSM	Reaction coefficients/Switching threshold
Frijns et al. (2010)	ABS	EMS	Local volatility/Belief coefficients/Intensity of choice
Franke and Westerhoff (2011)	IAH	MSM	Behavioural rules/Flexibility/Predisposition coefficients
Franke and Westerhoff (2012)	ABS/IAH	MSM	Behavioural rules/Wealth/Predisposition/Misalignment coefficients
Ghoshghadze and Lux (2015)	IAH	GMM	Standard deviation of innovations/Herding tendency
Grazzini et al. (2013)	Bass (1969)	ML, MSM	Probability of independent adoption/Peer pressure/Population size
Grazzini and Richiardi (2015)	-	SMD	-
Goldbaum and Zwinkels (2014)	?	OLS (iterative)	Belief coefficients
Hommes and Veld (2015)	ABS	NLS(?)	Belief coefficients/A-synchronous updating ratio(?)/Intensity of choice
Huisman et al. (2010)	ABS	Quasi ML	Belief coefficients/Intensity of choice
Kouwenberg and Zwinkels (2014)	ABS	Quasi ML	Belief coefficients/Intensity of choice
Kouwenberg and Zwinkels (2015)	ABS	Quasi ML	Price elasticity/Belief coefficients/Intensity of choice
Lof (2012)	ABS	NLS	Belief coefficients/Intensity of choice
Lof (0)	ABS	VAR/NLS	Discount factors/Belief coefficients/Intensity of choice
Reitz and Slopek (2009)	ABS	Quasi ML	GARCH coefficients/Belief coefficients/Transition parameter
Recchioni et al. (2015)	ABS	calibration	Belief coefficient/Intensity of choice/Risk aversion/Fundamental value/Memory
Verschoor and Zwinkels (2013)	ABS	ML	Belief coefficients/Intensity of choice

*Note:* The Table follows the logic of Table 1 and summarises recent research not covered there. Authors are alphabetised. The full meaning of the acronyms under ‘Origin’: Cusp stands for the cusp catastrophe model, ABS for Adaptive Belief System, and IAH for Interactive Agent Hypothesis. The full meaning of the acronyms under ‘Methods’: ML stands for Maximum Likelihood, RV for Realised Volatility, NLS for Nonlinear Least Squares, VAR for Vector Autoregression, MSM (SMM) for Method of Simulated Moments, OLS for Ordinary Least Squares, EMS for Empirical Martingale Simulation by Duan and Simonato (1998), GMM for Generalized Method of Moments, and SMD for Simulated Minimum Distance. ‘?’ means that given information is unclear to authors.



standard deviations of estimates preventing from contributive interpretation of results. In the majority, especially older, studies a discussion about the shape of the log-likelihood function is missing and the reader might only guess from insignificant estimates of the switching coefficient. A few most recent studies report that the likelihood is not very informative and the model accuracy is not sensitive for given parameter, and “the other parameters can to a large extent compensate for changes in  $\beta$ ”, the switching coefficient (Bolt et al., 2014, pg. 15). However, the shape of the objective function is almost never rigorously studied. An exception is e.g. Hommes and Veld (2015), who emphasise a very flat shape of the likelihood function for the intensity of choice selection that hampers validity of the test to reject the null hypothesis of switching, especially for small samples. On the other hand, smoothness of the objective function does not seem to be an issue for ML methods—this finding is further confirmed through Sections 4 and 5 also for Non-Parametric Simulated Maximum Likelihood (NPSML).

### 2.3. *Switching*

Finally, a high importance is devoted to the existence of behavioural switching, that is, to the sign, magnitude, and the statistical significance of the intensity of choice. Following the question of Chen et al. (2012, pg. 202), “how big or how small is it?”, we, however, need to emphasise that the magnitude of the switching coefficient cannot be directly compared across various models, assets, or time periods, as it is a unit free variable and its effect on the model dynamics is conditional on the particular model design and data. Four studies find a very large switching coefficients (Bolt et al., 2011, 2014; ter Ellen et al., 2013; Frijns et al., 2010), however, statistically insignificant in all cases. In other relevant studies (20 out of 23), the estimated values are mostly found single-digit and often close to zero, that well resembles the economic intuition of some, but realistically low switching frequency between major types of trading strategies. Although the sign of the parameter is of a crucial importance, in Tables 1 and 9 in Appendix 1 we present absolute values because the interpretation of the positive/negative sign also depends on the specific design of the model. Almost all studies report the theoretically expected sign of the effect, nonetheless, we do not observe any conclusive results regarding the statistical significance of the intensity of choice—no connection can be observed w.r.t. the ‘#’ number of estimated parameters or the frequency and length of the data. Statistically significant and insignificant findings are reported across these categories without any clear pattern. On the other hand, some but definitely hypothetical relation might be observed based on the ‘Type’ of the data: statistically significant estimates strongly dominate for commodities and weakly prevail for stock markets; insignificant estimates prevail for real estate markets and dominate for FX. However, as the sample of studies is rather small and often problematically mutually comparable, these findings should be interpreted with a high caution.



### 3. Simulation-based estimation of FABMs: the case of Brock & Hommes HAM

This section describes an innovative general computational framework for empirical estimation of full-fledged FABMs that often lack the closed-form solution of the likelihood function. Some few authors thus apply simulation-based estimation MMs (for details see [Section 2](#)). We follow the [Kristensen and Shin \(2012\)](#) concept of simulated MLE based on nonparametric kernel methods. The methodology has been developed for dynamic models where no closed-form representation of the likelihood function exists and thus we cannot derive the usual MLE. Therefore it constitutes an opportune estimation method for general class of FABMs.

In [Sections 4](#) and [5](#) we adopt the NPSMLE method to the FABM literature and test its capability on the most famous and widely analysed model developed by [Brock and Hommes \(1998\)](#) for which we customise the general framework of [Kristensen and Shin \(2012\)](#). [Section 2](#) summarises other attempts to estimate models derived from [Brock and Hommes \(1998\)](#) approach that builds on evolutionary switching between trading strategies.

#### 3.1. The [Brock and Hommes \(1998\)](#) model

Our modelling framework is within the [Brock and Hommes \(1998\)](#) HAM. The model is a financial market application of the ABS—the endogenous, evolutionary selection of heterogeneous expectation rules following the framework of [Lucas \(1978\)](#) and proposed in [Brock and Hommes \(1997, 1998\)](#). We consider an asset pricing model with one risk free and one risky asset. The dynamics of the wealth is as follows:

$$W_{t+1} = RW_t + (p_{t+1} + y_{t+1} - Rp_t)z_t, \quad (1)$$

where  $W_{t+1}$  stands for the total wealth at time  $t+1$ ,  $p_t$  denotes the ex-dividend price per share of the risky asset at time  $t$ , and  $\{y_t\}$  denotes its stochastic dividend process. The risk-free asset is perfectly elastically supplied at constant gross interest rate  $R = 1 + r$ , where  $r$  is the interest rate. Finally,  $z_t$  denotes the number of shares of the risky asset purchased at time  $t$ . The type of utility function considered is essential for each economic model and determines its nature and dynamics. The utility for each<sup>5</sup> investor (trader or agent alternatively)  $h$  is given by  $U(W) = -\exp(-aW)$ , where  $a > 0$  denotes the risk aversion, which is assumed to be equal for all investors. For determining the market prices in this model, the Walrasian auction scenario is assumed. I.e. the market clearing price  $p_t$  is defined as the price that makes demand for the risky asset equal to supply at each trading period  $t$  and investors are ‘price takers’. The detailed description of the price formation mechanism is offered further in this section and finally summarised by [Equation 13](#) and [Equation 14](#).

Let  $E_t$ ,  $V_t$  denote the conditional expectation and conditional variance operators, respectively, based on a publicly available information consisting of past prices and dividends, i.e. on the information set  $\mathcal{F}_t = \{p_t, p_{t-1}, \dots; y_t, y_{t-1}, \dots\}$ . Let  $E_{h,t}$ ,  $V_{h,t}$  denote the beliefs of investor type  $h$  (trader type  $h$  alternatively) about the conditional expectation and conditional variance. For analytical tractability, beliefs about the conditional variance of excess returns are assumed to be constant and the same for all investor types, i.e.  $V_{h,t}(p_{t+1} + y_{t+1} - Rp_t) = \sigma^2$ . Thus the conditional variance of total wealth  $V_{h,t}(W_{t+1}) = z_t^2 \sigma^2$ .

---

<sup>5</sup>This is a crucial assumption without which the original model of [Brock and Hommes \(1998\)](#) loses one of its greatest advantages of analytical tractability.

Each investor is assumed to be a myopic<sup>6</sup> mean variance maximiser, so for each investor  $h$  the demand for the risky asset  $z_{h,t}$  is the solution of:

$$\max_{z_t} \left\{ E_{h,t}[W_{t+1}] - \frac{a}{2} V_{h,t}[W_{t+1}] \right\}. \quad (2)$$

Thus

$$E_{h,t}[p_{t+1} + y_{t+1} - Rp_t] - a\sigma^2 z_{h,t} = 0, \quad (3)$$

$$z_{h,t} = \frac{E_{h,t}[p_{t+1} + y_{t+1} - Rp_t]}{a\sigma^2}. \quad (4)$$

Let  $n_{h,t}$  be the fraction of investors of type  $h$  at time  $t$  and its sum is one, i.e.  $\sum_{h=1}^H n_{h,t} = 1$ . Let  $z_{s,t}$  be the overall supply of outside risky shares. The Walrasian temporary market equilibrium for demand and supply of the risky asset then yields:

$$\sum_{h=1}^H n_{h,t} z_{h,t} = \sum_{h=1}^H n_{h,t} \left\{ \frac{E_{h,t}[p_{t+1} + y_{t+1} - Rp_t]}{a\sigma^2} \right\} = z_{s,t}, \quad (5)$$

where  $H$  is the number of different investor types. In the simple case  $H = 1$  we obtain the equilibrium pricing equation and for the specific case of zero supply of outside risky shares, i.e.  $z_{s,t} = 0$  for all  $t$ , the market equilibrium then satisfies:

$$Rp_t = \sum_{h=1}^H n_{h,t} \{E_{h,t}[p_{t+1} + y_{t+1}]\}. \quad (6)$$

In a completely rational market Equation 6 reduces to  $Rp_t = E_t[p_{t+1} + y_{t+1}]$  and the price of the risky asset is completely determined by economic fundamentals and given by the discounted sum of its future dividend cash flow:

$$p_t^* = \sum_{k=1}^{\infty} \frac{E_t[y_{t+k}]}{(1+r)^k}, \quad (7)$$

where  $p_t^*$  depends upon the stochastic dividend process  $\{y_t\}$  and denotes the fundamental price which serves as a benchmark for asset valuation based on economic fundamentals under rational expectations. In the specific case where the process  $\{y_t\}$  is independent and identically distributed,  $E_t\{y_{t+1}\} = \bar{y}$  is a constant. The fundamental price, which all investors are able to derive, is then given by the simple formula:

$$p^* = \sum_{k=1}^{\infty} \frac{\bar{y}}{(1+r)^k} = \frac{\bar{y}}{r}. \quad (8)$$

For the further analysis it is convenient to work not with the price levels, but with the deviation  $x_t$  from the fundamental price  $p_t^*$ :

$$x_t = p_t - p_t^*. \quad (9)$$

---

<sup>6</sup>To be ‘myopic’ means to have a lack of long run perspective in planning. Roughly speaking, it is the opposite expression to ‘intertemporal’ in economic modelling.

### 3.1.1. Heterogeneous beliefs

Now we introduce the heterogeneous beliefs about future prices. We follow the [Brock and Hommes \(1998\)](#) approach and assume the beliefs of individual trader types in the form:

$$E_{h,t}(p_{t+1} + y_{t+1}) = E_t(p_{t+1}^* + y_{t+1}) + f_h(x_{t-1}, \dots, x_{t-L}), \quad \text{for all } h, t, \quad (10)$$

where  $p_{t+1}^*$  denotes the fundamental price ([Equation 7](#)),  $E_t(p_{t+1}^* + y_{t+1})$  denotes the conditional expectation of the fundamental price based on the information set  $\mathcal{F}_t = \{p_t, p_{t-1}, \dots; y_t, y_{t-1}, \dots\}$ ,  $x_t = p_t - p_t^*$  is the deviation from the fundamental price ([Equation 9](#)),  $f_h$  is some deterministic function which can differ across trader types  $h$  and represents a ‘ $h$ -type’ model of the market, and  $L$  denotes the number of lags.

It is now important to be very precise about the class of beliefs. From the expression in [Equation 10](#) it follows that beliefs about future dividends flow:

$$E_{h,t}(y_{t+1}) = E_t(y_{t+1}), \quad h = 1, \dots, H, \quad (11)$$

are the same for all trader types and equal to the true conditional expectation. In the case where the dividend process  $\{y_t\}$  is i.i.d., from [Equation 8](#) we know that all trader types are able to derive the same fundamental price  $p_t^*$ .

On the other hand, traders’ beliefs about future price abandon the idea of perfect rationality and move the model closer to the real world. The form of this class of beliefs:

$$E_{h,t}(p_{t+1}) = E_t(p_{t+1}^*) + f_h(x_{t-1}, \dots, x_{t-L}), \quad \text{for all } h, t, \quad (12)$$

allows prices to deviate from their fundamental value  $p_t^*$ , which is a crucial step in heterogeneous agent modelling.  $f_h$  allows individual trader types to believe that the market price will differ from its fundamental value  $p_t^*$ .

An important consequence of the assumptions above is that heterogeneous market equilibrium from [Equation 6](#) can be reformulated in the deviations form, which can be conveniently used in empirical and experimental testing. We thus use [Equation 9, 10](#) and the fact that  $\sum_{h=1}^H n_{h,t} = 1$  to obtain:

$$Rx_t = \sum_{h=1}^H n_{h,t} E_{h,t}[x_{t+1}] = \sum_{h=1}^H n_{h,t} f_h(x_{t-1}, \dots, x_{t-L}) \equiv \sum_{h=1}^H n_{h,t} f_{h,t}, \quad (13)$$

where  $n_{h,t}$  is the value related to the beginning of period  $t$ , before the equilibrium price deviation  $x_t$  has been observed. The actual market clearing price  $p_t$  might then be calculated simply using [Equation 9](#) as  $p_t = x_t + p_t^*$ , expressed more precisely, combining [Equation 8](#), [Equation 9](#), and [Equation 13](#) as:

$$p_t = x_t + p_t^* = \frac{\sum_{h=1}^H n_{h,t} f_{h,t}}{R} + \frac{\bar{y}}{r}. \quad (14)$$

### 3.1.2. Selection of strategies

Beliefs of individual trader types are updated evolutionary and thus create the ABS, where the selection is controlled by endogenous market forces ([Brock and Hommes, 1997](#)). It is actually an expectation feedback system as variables depend partly on the present values and partly on the future expectations.

The profitability (performance) measures for strategies  $h$ ,  $h = 1, \dots, H$  are derived from past

realised profits as:<sup>7</sup>

$$U_{h,t} = (x_t - Rx_{t-1}) \frac{f_{h,t-1} - Rx_{t-1}}{a\sigma^2}. \quad (15)$$

Market fractions of trader types  $n_{h,t}$  are then given by the discrete choice probability—the multinomial logit model:<sup>8</sup>

$$n_{h,t} = \frac{\exp(\beta U_{h,t-1})}{Z_t}, \quad (16)$$

$$Z_t \equiv \sum_{h=1}^H \exp(\beta U_{h,t-1}), \quad (17)$$

where the one-period-lagged timing of  $U_{h,t-1}$  ensures that all information for the market fraction  $n_{h,t}$  updating is available at the beginning of period  $t$ ,  $\beta$  is the intensity of choice parameter measuring how fast traders are willing to switch between different strategies.  $Z_t$  is then normalisation ensuring  $\sum_{h=1}^H n_{h,t} = 1$ .

### 3.1.3. Basic belief types

In the original paper by [Brock and Hommes \(1998\)](#), the authors analyse the behaviour of the artificial market consisting of a few simple belief types (trader types or strategies). The aim of investigating the model with only two, three, or four belief types is to describe the role of each particular belief type in deviation from fundamental price and to investigate the complexity of the simple model dynamics with the help of the bifurcation theory.

All beliefs have the simple linear form:

$$f_{h,t} = g_h x_{t-1} + b_h, \quad (18)$$

where  $g_h$  denotes the trend parameter and  $b_h$  is the bias of trader type  $h$ . This form comes from the argument that only very simple forecasting rules can have a real impact on equilibrium prices as complicated rules are unlikely to be learned and followed by sufficient number of traders. [Hommes \(2006\)](#) also notices another important feature of [Equation 18](#), which is that  $x_{t-1}$  is used to forecast  $x_{t+1}$ , because [Equation 5](#) has not revealed equilibrium  $p_t$  yet when  $p_{t+1}$  forecast is estimated.

The first belief type are fundamentalists or rational ‘smart money’ traders. They believe that the asset price is determined solely by economic fundamentals according to the EMH introduced in [Fama \(1970\)](#) and computed as the present value of the discounted future dividends flow. Fundamentalists believe that prices always converge to their fundamental values. In the model, fundamentalists comprise the special case of [Equation 18](#) where  $g_h = b_h = f_{h,t} = 0$ . It is important

---

<sup>7</sup>Additional memory can be introduced into the profitability measure ([Equation 15](#)) e.g. as a weighted average of past realised values  $U_{m,h,t} = U_{h,t} + \eta U_{m,h,t-1}$ , where  $0 \leq \eta \leq 1$  denotes the ‘dilution parameter’ of the past memory in the profitability measure. Nonetheless, for the majority of examples, [Brock and Hommes \(1998\)](#) use  $\eta = 0$  to keep derivations analytically tractable and work with models without memory, i.e. [Equation 15](#) specification is used directly.

<sup>8</sup>With regard to Macro ABMs, [Branch and Evans \(2006, pg. 266\)](#) point out that “the multinomial logit has proven to be an important approach to modelling economic choices, and has been increasingly employed in recent work in dynamic macroeconomics”.

to note that fundamentalists' demand also reflects market actions of other trader types. Fundamentalists have all past market prices and dividends in their information set  $\mathcal{F}_{h,t}$ , but they are not aware of the fractions  $n_{h,t}$  of other trader types. Fundamentalists might pay costs  $C \geq 0$  to learn how fundamentals work and to obtain market information. However, [Brock and Hommes \(1998\)](#) themselves mostly set  $C = 0$  to keep simplicity of the analysis.

Chartists or technical analysts, sometimes called 'noise traders' represent another belief type. They believe that asset price is not determined by economic fundamentals only, but it can be partially predicted using simple technical trading rules, extrapolation techniques or taking various patterns observed in the past prices into account. If  $b_h = 0$ , trader  $h$  is called a pure trend chaser if  $0 < g_h \leq R$  and a strong trend chaser if  $g_h > R$ . Additionally, if  $-R \leq g_h < 0$ , the trader  $h$  is called contrarian or strong contrarian if  $g_h < -R$ .

Next, if  $g_h = 0$  trader  $h$  is considered to be purely upward biased if  $b_h > 0$  or purely downward biased if  $b_h < 0$ .

### 3.2. Construction of the NPSMLE

This subsection introduces the estimation framework for the [Brock and Hommes \(1998\)](#) model. Let us assume processes  $(x, v)$ ,  $x : t \mapsto \mathbb{R}^k$ ,  $v : t \mapsto \mathcal{V}_t$ ,  $t = 1, \dots, \infty$ . The space  $\mathcal{V}_t$  can be time-varying. Suppose that we have  $T$  realisations  $\{(x_t, v_t)\}_{t=1}^T$ . Let us further assume the time series  $\{x_t\}_{t=1}^T$  has been generated by a fully parametric model:

$$x_t = q_t(v_t, \varepsilon_t, \theta), \quad t = 1, \dots, T, \quad (19)$$

where a function  $q : \{v_t, \varepsilon_t, \theta\} \mapsto \mathbb{R}^k$ ,  $\theta \in \Theta \subseteq \mathbb{R}^l$  is an unknown parameter vector, and  $\varepsilon_t$  is an independent identically distributed (i.i.d.) sequence with known distribution  $\mathcal{F}_\varepsilon$ , which is (without loss of generality) assumed not to depend on  $t$  or  $\theta$ . In general, the processes  $(x, v)$  can be non-stationary and  $v_t$  is allowed to contain other exogenous variables than lagged  $x_t$ . We also assume the model to have an associated conditional density  $c_t(x|v; \theta)$ , i.e.

$$C(x \in A | v_t = v) = \int_A c_t(x|v; \theta) dx, \quad t = 1, \dots, T, \quad (20)$$

for any Borel set  $A \subseteq \mathbb{R}^k$ .

Let us now suppose that  $c_t(x|v; \theta)$  from [Equation 20](#) does not have a closed-form representation. In such situation, we are not able to derive the exact likelihood function of the model in [Equation 19](#) and thus a natural estimator of  $\theta$ —the maximiser of the conditional log-likelihood:

$$\tilde{\theta} = \arg \max_{\theta \in \Theta} L_T(\theta), \quad L_T(\theta) = \sum_{t=1}^T \log c_t(x_t|v_t; \theta) \quad (21)$$

is not feasible.

In such situation, however, we are still able to simulate observations from the model in [Equation 19](#) numerically.<sup>9</sup> The method presented allows us to compute a simulated conditional density, which we further use to gain a simulated version of the MLE.

---

<sup>9</sup>For cases in which the model in [Equation 19](#) is itself intractable and thus we cannot generate observations from the exact model, [Kristensen and Shin \(2012\)](#) suggest a methodology for approximate simulations and define regularity conditions for the associated approximate NPSMLE  $\hat{\theta}_M$  to have the same asymptotic properties as the simulated MLE  $\hat{\theta}$  defined in [Equation 25](#).

To obtain a simulated version of  $c_t(x_t|v_t; \theta) \forall t \in \langle 1, \dots, T \rangle, x \in \mathbb{R}^k, v \in \mathcal{V}_t$ , and  $\theta \in \Theta$ , we firstly generate  $N \in \mathbb{N}$  i.i.d. draws from  $\mathcal{F}_\varepsilon$ ,  $\{\varepsilon_i\}_{i=1}^N$ , which are used to compute:

$$X_{t,i}^\theta = q_t(v_t, \varepsilon_i, \theta), \quad i = 1, \dots, N. \quad (22)$$

These  $N$  simulated i.i.d. random variables,  $\{X_{t,i}^\theta\}_{i=1}^N$ , follow the target distribution by construction:  $X_{t,i}^\theta \sim c_t(\cdot|v_t; \theta)$ , and therefore can be used to estimate the conditional density  $c_t(x|v; \theta)$  with kernel methods—we define:

$$\hat{c}_t(x_t|v_t; \theta) = \frac{1}{N} \sum_{i=1}^N K_\eta(X_{t,i}^\theta - x_t), \quad (23)$$

where  $K_\eta(\psi) = K(\psi/\eta)/\eta^k$ ,  $K: \mathbb{R}^k \mapsto \mathbb{R}$  is a generic kernel and  $\eta > 0$  is a bandwidth. Under regularity conditions on  $c_t$  and  $K$ , we get:

$$\hat{c}_t(x_t|v_t; \theta) = c_t(x_t|v_t; \theta) + O_P(1/\sqrt{N\eta^k}) + O_P(\eta^2), \quad N \longrightarrow \infty, \quad (24)$$

where the last two terms are  $o_P(1)$  if  $\eta \longrightarrow 0$  and  $N\eta^k \longrightarrow \infty$ .

Having obtained the simulated conditional density  $\hat{c}_t(x_t|v_t; \theta)$  from [Equation 23](#), we can now derive the simulated MLE of  $\theta$ :

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \hat{L}_T(\theta), \quad \hat{L}_T(\theta) = \sum_{t=1}^T \log \hat{c}_t(x_t|v_t; \theta). \quad (25)$$

The same draws are used for all values of  $\theta$  and we may also use the same set of draws from  $\mathcal{F}_\varepsilon(\cdot)$ ,  $\{\varepsilon_i\}_i^N$ , across  $t$ . Numerical optimization is facilitated if  $\hat{L}_T(\theta)$  is continuous and differentiable in  $\theta$ . Considering [Equation 23](#), if  $K$  and  $\theta \mapsto q_t(v, \varepsilon, \theta)$  are  $s \geq 0$  continuously differentiable, the same holds for  $\hat{L}_T(\theta)$ .

Under the regularity condition, the fact that  $\hat{c}_t(x_t|v_t; \theta) \xrightarrow{P} c_t(x_t|v_t; \theta)$  implies that also  $\hat{L}_T(\theta) \xrightarrow{P} L_T(\theta)$  as  $N \longrightarrow \infty$  for a given  $T \geq 1$ . Thus, the simulated MLE,  $\hat{\theta}$ , retains the same properties as the infeasible MLE,  $\tilde{\theta}$ , as  $T, N \longrightarrow \infty$  under suitable conditions.

### 3.3. Advantages and disadvantages

To quote from [Kristensen and Shin \(2012, pg. 85\)](#), “one of the merits of NPSML is its general applicability”. Authors also provide three examples of application of the methodology in their article. The first comprises an estimation of the short-term interest rate model of [Cox et al. \(1985\)](#). The second applies the methodology to a jump-diffusion model of daily S&P500 returns by [Andersen et al. \(2002\)](#). In the third example the general capabilities of the NPSMLE to estimate a generic Markov decision processes are examined.

[Kristensen and Shin \(2012\)](#) also report several advantages and disadvantages of the proposed estimator. Starting with the former, the estimator works whether the observations  $x_t$  are i.i.d. or non-stationary because the density estimator based on i.i.d. draws is not affected by the dependence structures in the observed data. Second, the estimator does not suffer from the curse of dimensionality, which is usually associated with kernel estimators. In general, high dimensional models, i.e. with larger  $k \equiv \dim(x_t)$  as we smooth only over  $x_t$  here, require larger number of simulations to control the variance component of the resulting estimator. However, the summation

in Equation 25 reveals an additional smoothing effect and the additional variance of  $\hat{L}_T(\theta)$  caused by simulations retains the standard parametric rate  $1/N$ .

Conversely, the simulated log-likelihood function is a biased estimate of the actual log-likelihood function for fixed  $N$  and  $\eta > 0$ . To obtain consistency, we need  $N \rightarrow \infty$  and  $\eta \rightarrow 0$ . Thus, the parameter  $\eta$  needs to be properly chosen for given sample and simulation size. In the stationary case, the standard identification assumption is:

$$\mathbb{E}[\log c(x_t|v_t, \theta)] < \mathbb{E}[\log c(x_t|v_t, \theta_0)] \quad \forall \theta \neq \theta_0. \quad (26)$$

Under stronger identification assumptions, the choice of the parameter  $\eta$  might be less important and one can prove the consistency of the estimator for any fixed  $0 < \eta < \bar{\eta}$  for some  $\bar{\eta} > 0$  as  $N \rightarrow \infty$  (Altissimo and Mele, 2009). In practice this still requires us to know the threshold level  $\bar{\eta} > 0$  but from the theoretical viewpoint this ensures that parameters can be well identified in large finite samples after a given  $\bar{\eta} > 0$  is set. Moreover, it suggests that proposed methodology is fairly robust to the choice of  $\eta$ . In their simulation study, Kristensen and Shin (2012) show indeed that the NPSMLE performs well using broad range of bandwidth choices.

### 3.4. Asymptotic properties

As the theoretical convergence of the simulated conditional density towards the true density is met, we would expect the NPSMLE  $\hat{\theta}$  to have the same asymptotic properties as the infeasible MLE  $\tilde{\theta}$  for a properly chosen sequence  $N = N(T)$  and  $\eta = \eta(N)$ . Kristensen and Shin (2012) show that  $\hat{\theta}$  is first-order asymptotic equivalent to  $\tilde{\theta}$  under set a general conditions, allowing even for non-stationary and mixed discrete and continuous distribution of the response variable. Further, using additional assumptions, including stationarity, they provide results regarding the higher-order asymptotic properties of  $\hat{\theta}$  and derive expressions of the bias and variance components of the NPSMLE  $\hat{\theta}$  compared to the actual MLE due to kernel approximation and simulations.

Therefore, a set of general conditions, satisfied by most models, need to be verified so that  $\hat{c} \rightarrow c$  sufficiently fast to ensure asymptotic equivalence of  $\hat{\theta}$  and  $\tilde{\theta}$ . Kristensen and Shin (2012) define a set of regularity conditions on the model and its associated conditional density that satisfy these general conditions for uniform rates of kernel estimators defined in Kristensen (2009).

The kernel  $K$  from Equation 23 has to belong to a broad class of so-called bias high-order or bias reducing kernels. E.g. the Gaussian kernel, which we use in Section 4, satisfy this condition if  $r \geq 2$ , where  $r$  is the number of derivatives of  $c$ . Higher  $r$  causes faster rate of convergence and determines the degree of  $\hat{c}$  bias reduction. Moreover, general versions of conditions usually required for consistency and well-defined asymptotic distribution (asymptotic normality) of MLEs in stationary and ergodic models are imposed on actual log-likelihood function and the associated MLE to ensure the actual MLE  $\tilde{\theta}$  in Equation 21 is asymptotically well-behaved.

## 4. Monte Carlo study: NPSMLE of the HAM

This section analyses the capability of the NPSMLE methodology for the HAMs estimation purposes and evaluates small sample properties of the estimator via an extensive Monte Carlo study. We simulate data from the HAM and employ the NPSMLE of selected model parameters to analyse how well and under what conditions is the estimation method able to recover true values of parameters in the controlled environment. For its conceptual importance, a detailed focus is devoted to the switching parameter—intensity of choice  $\beta$ .



#### 4.1. Simulation setup for the HAM

In the simulation setup, we follow the previous works of [Barunik et al. \(2009\)](#); [Vacha et al. \(2012\)](#); [Kukacka and Barunik \(2013\)](#). The joint setup for the basic HAM model (see [Subsection 3.1](#)) is used for all (if not explicitly stated otherwise) conducted simulations in this section and is defined as follows. The model we use to generate observations is a very stylised simple version compactly described in [Hommes \(2006, pg. 1169\)](#) and consisting of three mutually dependent equations:

$$Rx_t = \sum_{h=1}^H n_{h,t} f_{h,t} + \epsilon_t \equiv \sum_{h=1}^H n_{h,t} (g_h x_{t-1} + b_h) + \epsilon_t, \quad (27)$$

$$n_{h,t} = \frac{\exp(\beta U_{h,t-1})}{\sum_{h=1}^H \exp(\beta U_{h,t-1})}, \quad (28)$$

$$\begin{aligned} U_{h,t-1} &= (x_{t-1} - Rx_{t-2}) \frac{f_{h,t-2} - Rx_{t-2}}{a\sigma^2} \\ &\equiv (x_{t-1} - Rx_{t-2}) \frac{g_h x_{t-3} + b_h - Rx_{t-2}}{a\sigma^2}, \end{aligned} \quad (29)$$

where  $\epsilon_t$  (which coincides with  $\varepsilon_t$  in [Equation 19](#)) is an i.i.d. noise term sequence with given distribution<sup>10</sup> representing the market uncertainty and unpredictable market events.

In order to run the model in various different settings, we inevitably need to fix several variables less important for the dynamics of the model to enable estimation of the key parameters. First, we set the constant gross interest rate  $R = 1 + r = 1.0001$  to resemble real market risk free rate. Assuming 250 trading days per year and daily compounding, this daily value represents circa 2.5% annual risk free interest rate which is a reasonable approximation. Although this figure is not based on any rigorous calibration or taken from a specific study, similar values are largely used in various financial and macroeconomic works. Moreover, as we show in further analysis, the model exhibits considerable robustness w.r.t. various reasonable risk free values and thus there is no need for more precise derivation of this parameter. We further fix the linear term  $1/a\sigma^2$  (comprising the risk aversion coefficient  $a > 0$  and the beliefs about the conditional variance of excess returns  $\sigma^2$ ) to 1. The similar setting has already been successfully used in previous works of [Barunik et al. \(2009\)](#); [Vacha et al. \(2012\)](#); [Kukacka and Barunik \(2013\)](#). It is important to note that  $a$  and  $\sigma^2$  are only scale factors for the profitability measure  $U$ . Their magnitudes do not affect relative proportions of  $U_{h,t}$  and thus do not influence the dynamics of the model output, that is on the contrary usually characterised by time-varying variance. In other words, although we assume constant  $\sigma^2$ , the output time series generated by the model does not have constant variance. Strategy-specific  $a_h$  or time-varying  $\sigma_{h,t}^2$  are appealing concepts mainly for simulation analyses of HAMs (see e.g. [Gaunersdorfer, 2000](#); [Chiarella and He, 2002](#); [Amilon, 2008](#)). Moreover, we intentionally use relatively small number of possible trading strategies following ([Kukacka and Barunik, 2013](#)),  $H = 5$ , for the general model setting or  $H \in \{2, 3\}$  for so called 2-type and 3-type model, respectively ([Chen et al., 2012](#)). Following [Hommes \(2006\)](#) via [Equation 29](#), neither ‘dilution parameter’ of the past memory  $\eta$  nor information costs  $C$  for fundamentalists are implemented into the basic model setup to keep the dynamics of the model and impacts of assessed modification as

---

<sup>10</sup>Various specifications of normal and uniform distributions are utilised in [Section 4](#), standard deviation of a normal distribution is estimated in [Section 5](#).



clear as possible. Indeed, also [Brock and Hommes \(1998\)](#) mostly set  $C = 0$  to keep simplicity of the analysis and work with models without memory, i.e. they set the  $\eta = 0$  (see [Subsection 3.1.2](#)) to keep derivations analytically tractable.

Within the Monte Carlo method, several parameters are repeatedly randomly generated to obtain statistically valid inference. Following the previous works by [Barunik et al. \(2009\)](#); [Vacha et al. \(2012\)](#); [Kukacka and Barunik \(2013\)](#), trend parameters  $g_h$  are drawn from the normal distribution  $N(0, 0.4^2)$  and bias parameters  $b_h$  are drawn from the normal distribution  $N(0, 0.3^2)$ . ‘Strict’ fundamental strategy in the sense of the original [Brock and Hommes \(1998, pg. 1245\)](#) article appears in the market by default, i.e. the first strategy is always defined as  $g_1 = b_1 = 0$  and therefore fundamentalists are always present on the market.

In the Monte Carlo simulations, we first study the capabilities of the NPSMLE under various levels of the switching parameter—intensity of choice  $\beta$ . As discussed in [Section 2](#), literature estimating  $\beta$  using real marked data is relatively scarce because of difficulties arising from the nonlinear nature of the HAM. Thus,  $\beta$  still remains a rather theoretical concept. Larger  $\beta$  implies higher willingness of agents to switch between available trading strategies based on their relative profitability—the best strategy attracts the most agents at each period. On the one hand, comprising the large variety of possible  $\beta$  values might seem as a dominant simulation strategy, on the other hand one has to consider computational burden of the simulation process in real time. What is perhaps even more important is to consider intensity of choice  $\beta$  from the economic viewpoint. First, high values give rise to unrealistically high switching frequency, which is hardly to be observed among market agents in reality. Next, negative  $\beta$  does not make any economic sense in the presented model framework as it causes inverse illogical switching towards less profitable strategies. Although the intensity of choice  $\beta$  cannot be directly rigorously compared across various models, assets, or time periods (see discussion in [Section 2](#)), we utilise the general knowledge of previous estimation efforts for models sharing similar framework to set meaningful simulation grids in this section. When referring to literature review (see [Subsection 2.3](#)), in vast majority of research articles sharing the ABS framework derived from [Brock and Hommes \(1998\)](#),  $\beta$  is found single-digit and often close to zero, that well resembles the economic intuition of some, but realistically low switching frequency between major types of trading strategies. Thus, we employ relatively rich, but reasonable discrete range of  $\beta$  in our simulations:  $\{0, 0.1, 0.5, 1, 3, 5, 10\}$ . It is far beyond the scope of this work to provide a deep analysis of the model behaviour, e.g. how the intensity of choice  $\beta$  influences the dynamics of the model that can under some setting even generate purely chaotic behaviour. Many studies have been devoted to this generally difficult issue in past two decades. In this context we refer the interested reader to the original paper of [Brock and Hommes \(1998\)](#) containing comprehensive model dynamics analysis, extensive studies by [Hommes \(2006\)](#), [Hommes and Wagener \(2009\)](#), [Chiarella et al. \(2009\)](#), or the recent book summarising 20 years of research on the Heterogeneous Expectations Hypothesis by [Hommes \(2013\)](#).

Next, as discussed by [Amilon \(2008\)](#), the magnitude of noise term has to be considered carefully. Noise is an inevitable part of the model as it represents the market uncertainty and unpredictable market events, but it must not overshadow the effect of variables under scrutiny. As mentioned in [Kukacka and Barunik \(2013\)](#), although varying noise variance can cause some minor changes in model outcomes, the analysed HAM embodies major similarities across various noise. Although theoretically the  $\mathcal{F}_\varepsilon$  from which the  $\{\varepsilon_i\}_{i=1}^N$  are drawn to simulate  $\{X_{t,i}^\theta\}_{i=1}^N$  ([Equation 22](#)) is a generic known distribution, the assumptions about market noise can play crucial role in the NPSMLE application to real world data. Therefore we test the model sensitivity and robustness

of proposed methodology using 30 stochastic noise specifications from an extensive range drawn a) from various normal distributions and b) from the uniform distributions

1. that cover the same intervals as are covered by their respective normal counterparts by the 99.74% of the probability mass;
2. with the same variances as their respective normal counterparts.

Detailed description of all 30 stochastic noise specifications can be found in [Table 10](#) (specification for normal distributions), [Table 13](#) (specification for uniform distributions of the 1. type), and [Table 14](#) (specification for uniform distributions of the 2. type). Basically, for the normally distributed noise the range extends from a ‘miniscule’ standard deviations  $SD = 10^{-8}, 10^{-7}, 10^{-6}$  [a value used by [Hommes \(2013, pg. 170, 174, 177\)](#) in a similar model setting], or  $10^{-5}$ , followed by ‘small’ standard deviations  $SD = 10^{-4}, 10^{-3}, 0.01$  [another value used by [Hommes \(2013, e.g. pg. 171\)](#) in a similar setting], standard normal  $SD = 1$ , and finally a relatively large ‘experimental’ standard deviation  $SD = 2$ . The sensitivity analysis of the NPSMLE method to the stochastic noise specification is based on the normality assumption. The normal distribution of market noise seems reasonably realistic and similar assumption has already been used in related studies, where “the non-linear models are fed with an exogenous stochastic process, but the noise process is ‘nice’, which in this case means that it is normally distributed”, as pointed out by [Amilon \(2008, pg. 344\)](#). We also utilise the favourable theoretical properties of the Gaussian kernel ([Kristensen and Shin, 2012, pg. 81](#)) in [Equation 23](#). To check the robustness of the method, we concur the previous research in [Barunik et al. \(2009\)](#); [Vacha et al. \(2012\)](#); [Kukacka and Barunik \(2013\)](#)—where uniform stochastic noise specification is utilised—and compare and contrast the results based on normally distributed noise to the two rather extreme and economically unrealistic uniform variants defined above. We intentionally do not consider any at first sight soliciting heavy-tailed noise distribution. The fact that financial data are heavy-tailed does not suggest any specific distribution of the market noise. In fact, the situation is opposite. The attractiveness of the HAM is based on its ability to produce heavy-tailed distribution of model output although we input normally distributed stochastic noise. Thus the HAM explains one of the most important stylised facts of financial time series via endogenous interactions of fundamentalists and boundedly rational chartists, not as an effect of a specific distribution of noise input. Finally, five lengths of the resulting series entering the NPSMLE algorithm are used: 100, 500, 1000, 5000, and 10000, and first 100 observations from the HAM are always discarded<sup>11</sup> as initial period, where the model dynamic is being established.

#### 4.2. Simulation setup for the NPSMLE

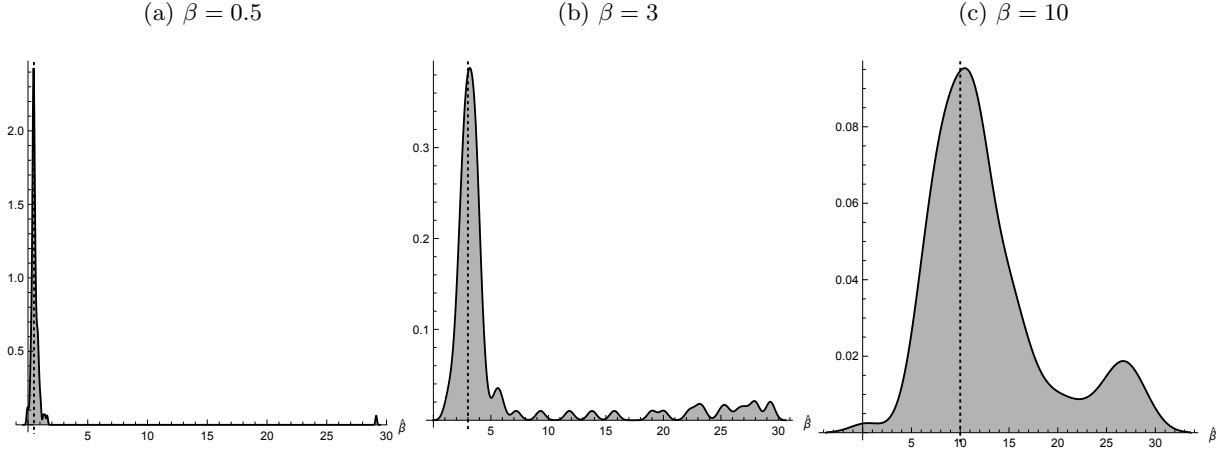
We follow the [Kristensen and Shin \(2012\)](#) methodology used for the estimation of the [Cox et al. \(1985\)](#) short-term interest rate model. However, we examine and adapt the setup for the purposes of the HAM. As discussed in [Subsection 3.3](#), there are two main trade-offs: between the precision of the kernel estimation and the computational burden, and between the smoothness of the kernel estimation and the bias.

As we are the first to apply a very recent NPSMLE methodology on a well-known HA modelling framework, we can only partially base our simulation setup on some results from literature. For this reason, we elaborate an extensive Monte Carlo simulation testing of the robustness of our

---

<sup>11</sup>We always simulate 100 extra observations to be discarded so that we finally get the intended length of the series with stable dynamics without the initialisation period.

Figure 1: Pre-estimation performance for selected  $\beta$ s



*Note:* Results based on 100 random runs, number of observations  $t = 1000$ , and the kernel estimation precision  $N = 100$ , initial point drawn from uniform distribution  $U(0, 30)$ . Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from standard normal distribution  $N(0, 1)$ . Black dotted vertical lines depict the true  $\beta$ s. Produced using automatic `SmoothHistogram` kernel approximation function in *Wolfram Mathematica*.

findings. First, to analyse statistically valid results, we start from the benchmark in [Kristensen and Shin \(2012\)](#) and compare the simulations of 100, 500, and 1000 runs. Moreover, three levels of the kernel estimation precision are considered, namely  $N = 100$ ,  $N = 500$ , and  $N = 1000$ . It is important to note that the same draws  $\{\varepsilon_i\}_{i=1}^N$  are used to generate the simulations  $\{X_{t,i}^\theta\}_{i=1}^N$  over time.

Second, the numerical algorithm is designed to find an optimum of either unconstrained or constrained multivariable function. As discussed in [Subsection 4.1](#), we expect  $\beta$  to be non-negative but rather small, i.e. single-digit. Using pre-estimation step with unconstrained parameter space we can obtain reasonably sufficient preliminary knowledge about the approximate true value of estimated parameters even for computationally feasible setting. This general principle of a preliminary rough search followed by a fine-tuning on a considerably restricted subset of the parameter space is successfully applied e.g. in [Chen and Lux \(2015\)](#) for MSM estimation. In [Figure 1](#) we demonstrate the pre-estimation performance via smooth histograms of  $\hat{\beta}$  based on a setting which can be easily computed using a personal computer within several minutes:  $\beta \in \{0.5, 3, 10\}$ , 100 runs, number of observations  $t = 1000$ , kernel estimation precision  $N = 100$ ,  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from standard normal distribution, and the initial point drawn from uniform distribution covering a broad interval  $\langle 0, 30 \rangle$ . We can observe how the peak of the distribution approximately detects the true value of  $\beta$  which helps us to constrain the parameter space in the next step. In subsequent—this time computationally very extensive—optimisation of the constrained function to fine-tune the precision of estimates we can therefore opt for relatively narrow bounds of the parameter space set as  $\langle -\beta, 3\beta \rangle$  for  $\beta \geq 0$  and  $\langle -0.5, 0.5 \rangle$  for  $\beta = 0$ .<sup>12</sup> For the 2-type model simulation estimation study

<sup>12</sup>This extended range of the parameter space bounds covering also economically irrelevant negative values is used to ensure the robustness of the method and not imposing excessive demands on the precision of the unconstrained pre-estimation. Moreover, not allowing for negative values might naturally lead to an upward bias of the simulated

(see [Subsection 4.3.6](#)), we use even wider and off-centered interval for the bounds of the parameter space set as  $\langle -3|g_2|, 3|g_2| \rangle$  and  $\langle -3|b_2|, 3|b_2| \rangle$ , respectively, to allow for possible negative values, and  $\langle -0.5, 0.5 \rangle$  for  $\beta = 0$ . For the 3-type model simulation estimation study (see [Subsection 4.3.10](#)), it is, however, important to limit bounds by zero from one side, i.e.  $\langle 0, 3|g_2| \rangle$ ,  $\langle -3|g_3|, 0 \rangle$  to avoid problems with insufficient specification of the model leading to ambiguous bimodal distributions of estimated parameters. The same intervals are used for a random draw of a single<sup>13</sup> starting point of the optimisation search procedure which is drawn from the uniform distributions.

Third, to estimate the conditional density  $c_t(x|v; \theta)$  with the kernel method ([Equation 23](#)), the Gaussian kernel and the [Silverman's \(1986\)](#) rule of thumb for finding the optimal size of the bandwidth:

$$\eta = \left( \frac{4}{3N} \right)^{1/5} \hat{\sigma}, \quad (30)$$

where  $\hat{\sigma}$  denotes the standard deviation of  $\{X_{t,i}^\theta\}_{i=1}^N$ , are employed.

Additionally, [Kristensen and Shin \(2012, pg. 82\)](#) suggest undersmoothing option for the bandwidth size selection concluding that “simulation results indicate that standard bandwidth selection rules together with a bit of undersmoothing in general deliver satisfactory results”. Moreover, as found by [Jones et al. \(1996\)](#), smaller bandwidths are better for larger kernel approximation precision, “because the estimator should be ‘more local’ when more information is present, and when the density is rougher, because the bias effect is stronger”. However, we do not use the undersmoothing option in our numerical algorithm as for HAM the methodology is robust in this aspect and various levels of undersmoothing do not change the outcomes.<sup>14</sup>

### 4.3. Monte Carlo results

In all simulations we are concerned in questions how accurately is the method able to recover the true values and how robust is the method with respect to various settings. For this reason, all tables (but not figures) in this section always report results based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$  i.i.d. draws from given distribution. Sample medians and means of the estimated values together with standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.<sup>15</sup>

#### 4.3.1. $\beta$ estimation in the general model

First and foremost, we assess the simplest but also the most crucial case of  $\beta$  estimation in the general model with  $H = 5$  possible trading strategies. The intensity of choice  $\beta$  is the most important parameter influencing the dynamics of the system through the multinomial logit model of a continuous adaptive evolution of market fractions in [Equation 16](#). Not only its magnitude between two extreme cases  $\beta = 0$  and  $\beta = \infty$  is important, but  $\beta$  also determines the type of

---

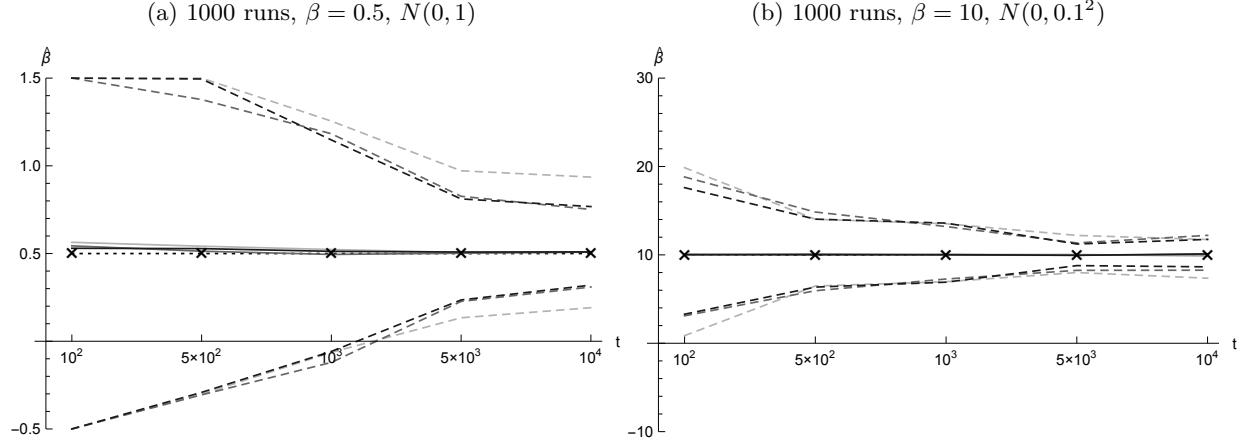
estimator especially for  $\beta$  close to 0.

<sup>13</sup>[Kristensen and Shin \(2012\)](#) use multiple starting points for the numerical optimisation but for the HAM estimation purposes in a simulated environment the single starting point is sufficient bringing the merits of markedly reduced computational time and burden.

<sup>14</sup>Results of this testing are available upon request from authors.

<sup>15</sup>We comment more on the issue of possible occurrence of ‘NaN’ outcome from the NPSMLE procedure in the following [Subsection 4.3.1](#).

Figure 2: Simulation results snapshot for  $\beta = 0.5; 10$



*Note:* Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distributions. Black dotted lines with  $\times$  depict the true  $\beta$ . Grey full lines depict sample means of estimated  $\beta$ . Grey dashed lines depict 2.5% and 97.5% quantiles. Light grey colour represents results for  $N = 100$ , normal grey for  $N = 500$ , and dark grey for  $N = 1000$ . ‘t’ (horizontal axis) stands for the length of generated time series.

the model equilibrium that can generally take the form of a (multiple) steady state(s), cycles, or even chaotic behaviour. The intensity of choice  $\beta$  is also crucial for its conceptual importance—it represents the dominant approach how the boundedly rational choices of agents are mathematically modelled in the current literature (see the ABS origin of models in Tables 1 and 2).

Despite of its relative simplicity, the setting is otherwise very challenging as capturing the effect of the switching coefficient  $\beta$  is generally difficult (see Section 2). Moreover, algorithm with a single starting point for the numerical optimisation and new random draws of the parameters  $g_h$  and  $b_h$ ,  $h \in \{2, 3, 4, 5\}$  for each independent run require very robust performance of the search procedure.

#### 4.3.2. Qualitative results

We primarily aim to verify whether important theoretical properties of the estimator, the consistency and asymptotic efficiency, also hold in small samples for the model. In Figures 2 and 7, 8 in Appendix 2 we depict and describe a ‘snapshot’ of simulation results for four interesting values of the intensity of choice  $\beta \in \{0, 0.5, 3, 10\}$  combined with three specifications of the stochastic noise:  $\epsilon_t \sim N(0, 2^2)$ ,  $\epsilon_t \sim N(0, 1)$  and  $\epsilon_t \sim N(0, 0.1^2)$ . First, we can clearly observe how the method is able to reveal the true value of  $\beta$  demonstrated by black dotted lines with  $\times$ . Grey full lines depict the sample means of estimated  $\beta$  and closely follow the true line. The small departures are naturally mainly observable for the smallest considered number of runs 100 [subfigures (a) and (b)] and for the smallest considered length of generated time series (number of observations)  $t = \{100, 500\}$ . Equally importantly, we can clearly observe the consistency of the estimator and how the efficiency of the mean estimate increases simultaneously with increasing length of generated time series  $t$  as well as the precision of the kernel estimation  $N$ . The precision is demonstrated by different shades of grey dashed lines depicting 2.5% and 97.5% quantiles of estimated parameters  $\beta$ . The shift from a relatively orderless pattern observed for 100 runs [subfigures (a) and (b)] to very exemplary

theoretically expected pattern with grey dashed lines nearly aligned according to increasing kernel estimation precision (from light grey farther and dark grey closer to the true/mean value) in (e) and (f) is obvious. The difference is much evident between  $N = 100$  (light grey) and two higher values, for  $N = \{500, 1000\}$  dashed lines are often very close to one another. The line representing  $N = 500$  actually appears closer to the mean value in some cases, indicating the sufficiency of the  $N = 500$  kernel approximation precision. Increasing statistical validity of results is apparent via rising number of runs starting at 100. Number of runs 500 seems sufficient in terms of only small differences compared to 1000 runs. In the right column of Figure 8 for  $\beta = 10$  the shift from the case of 100 runs (b) to the case of 1000 runs (f) is not so substantive as in the previous cases and we assign this to a more stable behaviour of the NPSMLE method under various settings caused by value of  $\beta$  farther from 0. But on the other hand, consistency of the estimator and the growth of efficiency of the mean estimate with increasing lengths of generated time series  $t$  and the precision of the kernel estimation  $N$  is well observable also in this setting.

Focusing further on the right column of Figure 7 and the left column of Figure 8, we moreover observe an important result from the economic interpretation point of view of the  $\beta$  parameter value. Although only the non-negative values of  $\beta$  have an economic interpretation,<sup>16</sup> in simulations we also allow for negative estimated values (for details of the simulation setup please refer to Subsection 4.1 and Subsection 4.2) to test the capability of the method even for such extreme values and to avoid the upward bias of the estimator. However, the most important result is that using a reasonably robust setting [e.g. number of runs 500 and precision of the kernel estimation  $N = 500$ ] we obtain more than 97.5% non-negative observation (represented by the bound of the 2.5% quantile to be found in the positive half-plane) for  $\beta = 10$  even when length of generated time series  $t = 100$ , for  $\beta = 3$  when  $t = 500$ , and for  $\beta = 0.5$  when  $t = 1000$ . At the same moment 95% of observation appear reasonably close to the true value, far from the numerical bounds of the parameter space imposed to make the constrained optimisation computationally feasible. These features have important favourable consequences for application of the method to datasets of various lengths—we should be able to detect even very weak signs of behavioural switching in long-span daily financial data, but also stronger signs of switching in macroeconomic data where typically lower-frequency time series of shorter lengths are available. W.r.t. the complexity of the estimation issue in the nonlinear HAM setting with five repeatedly randomly generated strategies (as well as to many other estimation attempts from Section 2 that have found the switching coefficient insignificant), we consider our results very promising. The most important property of the estimation method in the current setting is the ability to distinguish between statistically significant and insignificant  $\beta$  and this objective is well achieved.

Figures 7 and 8 also allow for comparison between estimation of models with and without switching. The left column of Figure 7 represents the model without switching ( $\beta = 0$ ), the right column and Figure 8 illustrate estimation performance for models with switching ( $\beta > 0$ ).

#### 4.3.3. Quantitative results

Now we move from the graphical to the quantitative description of simulation results. We now consider only the second most robust setting combination which proved optimal, i.e. results based on 1000 random runs, length of generated time series (number of observations)  $t = 5000$ , and the kernel estimation precision  $N = 1000$  i.i.d. draws from given distribution. First we comment

---

<sup>16</sup> $\beta < 0$  for which we technically allow implies switching of agent towards less profitable strategies, unambiguously economically irrational behaviour.



Table 3: Results for  $\beta$  estimation with normal noise

	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
$\beta$	(g) $\hat{\beta}$ , $N(0, 1)$						(h) $\hat{\beta}$ , $N(0, 0.1^2)$					
0	.00	.00	.11	-.24	.23	0%	.01	.01	.22	-.45	.46	0%
.1	.11	.11	.08	-.09	.30	0%	.11	.11	.12	-.10	.30	0%
.5	.50	.51	.14	.23	.81	0%	.50	.50	.35	-.35	1.30	0%
1	1.00	1.01	.23	.66	1.45	1%	.99	1.00	.50	-.19	2.46	0%
3	3.07	3.59	1.41	2.35	7.93	3%	2.99	3.05	1.00	1.48	5.86	0%
5	5.61	7.23	3.30	3.82	14.41	8%	4.99	5.05	1.21	3.75	6.81	1%
10	11.20	13.43	6.31	5.16	28.13	23%	9.99	9.99	2.22	7.57	11.64	3%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distributions of given parameters,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

on the robustness of the method w.r.t. various noise specifications used both for generating the stochastic term  $\epsilon_t$  in Equation 29 as well as for  $N$  i.i.d. draws,  $\{\varepsilon_i\}_{i=1}^N$ , to simulate  $N$  i.i.d. random variables,  $\{X_{t,i}^\theta\}_{i=1}^N$ , used for the kernel estimation of the conditional density. Again, following results offer a direct comparison between estimation of models with and without switching as first rows of all panels in all tables in Subsection 4.3.1 always represent the model setting without switching ( $\beta = 0$ ).

Interpreting results in Tables 3 and 10 in Appendix 2, the big picture seems promising for the NPSMLE method. We can observe relatively stable results over a reasonable grid of noise specifications and therefore the important issue of the robustness of the method is verified. Focusing on first columns containing the sample medians and means of the estimated values (denoted ‘Med.’ and ‘Mean’), we reveal the ability of the method to recover the true values of the intensity of choice  $\beta$  coefficient with very high precision over all noise specification. Median value is generally more precisely estimating the true value but the difference is negligible in majority of cases. Only for two most intensive noises in combination with higher values of the  $\beta$  coefficient, the mean estimate gives considerably better results. Comparing the third columns displaying related standard deviations we observe statistical significance of estimates for majority of combinations of the true  $\beta$  and the magnitude of noise. Generally the specifications with the smallest noises [subparts (a) and especially (b)] appear markedly more precise in estimating the lowest  $\beta s = \{0, 0.1, 0.5\}$  with only noise specification (b) having real ability estimate zero  $\beta$  with reasonable precision as we can observe using 2.5 (LQ) and 97.5 (HQ) quantile figures. On the other hand, the specifications with almost largest noises [subparts (i) and especially (h)] appear the most precise in estimating higher  $\beta s = \{1, 3, 5, 10\}$ . Values of the intensity of choice  $\beta$  very close to zero,  $\beta = \{0.1, 0.5\}$ , are the most difficult to estimate. This is, however, almost the extreme case of no switching of agents among possible strategies in which the dynamics of the model is restrained as there is only a small difference of the model behaviour compared to the agents’ absolute inertia case with  $\beta = 0$ . These are crucial findings highlighting the necessity of a proper noise specification within the estimation procedure. Larger noises seem to stabilise the system but overshadow the effect of switching under low  $\beta s$  and therefore favour estimation of higher  $\beta s$ . Lower  $\beta s$  require small noises for the effect of switching to be detectable. A puzzling result is then the subpart (h) with the largest noise intensity  $N(0, 2^2)$  estimating the lowest  $\beta s$  with high precision.

Another emerging issue is the occurrence of ‘Not a Number’ outcomes from some runs of the

NPSMLE procedure. This is reported via the NN column as the percentage of the ‘NaN’ outcomes. Technical reason behind the ‘NaN’ emergence is that the HAM algorithm does not converge into a stationary series in the particular run and the NPSMLE algorithm therefore produces a ‘NaN’ outcome. As we can see in Table 10, this situation is typical mainly for small distribution intervals of the stochastic noise  $\epsilon_t$  that do not always suffice to stabilise the system [see subparts (a) and (b)] and for highest values of the intensity of choice  $\beta$  [see subparts (i) and especially (j)] increasing switching dynamics in the model, particularly when these two effects combine together. We can interpret this as a specific kind of censorship of results as these runs are not considered for calculation of reported values. We do not consider results with high number of ‘NaN’ outcomes relevant within this analysis, however, we keep displaying them to retain the completeness of provided information as well as an optimal warning signal of an improper behaviour of the system under scrutiny. We can also observe signs of upward and downward biases of the estimates in cases of largest distribution intervals of the stochastic noise  $\epsilon_t$  when combined with highest  $\beta$ s [see subparts (i) and (j)] but this might just be an effect of this data censorship as all such cases are accompanied with occurrence of ‘NaN’ outcomes. Finally, one might notice that our grid of noise specifications in Table 10 to a great extent covers the entire range of reasonable values w.r.t. the ‘NaN’ emergence issue affecting estimation results crucially both for the smallest noises [see subparts (a) and (b)] as well as for the noise specifications with the largest standard deviation [subpart (j)]. Decreasing or increasing the noise intensity behind these bounds in specified setup leads to even less relevant results and therefore is not considered. As the ‘NaN’ emergence censoring estimation results might be a serious issue, it definitely needs more assessment in the next part of the study (see Subsection 4.3.5 and Subsection 4.3.6 for further discussion).

As we need to reduce the large grid of stochastic noise specification for other applications in Section 4, taking all discussed aspects into account, we select the two most ‘successful’ specifications, namely  $N(0, 0.1^2)$ —especially for larger  $\beta = \{3, 5, 10\}$ —and  $N(0, 1)$ —especially for smaller  $\beta$ s =  $\{0.5, 1\}$ .<sup>17</sup> They produce estimates with low standard errors for majority of  $\beta$ s considered, they are not accompanied with excessive number of ‘NaN’ outcomes, and they appear reasonably realistic w.r.t. the empirical application in Section 5 where we analyse time series of price deviations implying higher standard deviations of the assumed stochastic market noise. Figure 9 in Appendix 2 depicts smooth histograms of selected estimated  $\beta$ s based on these three noise specifications. One can clearly observe how the noise specification  $N(0, 0.1^2)$  performs best for  $\beta = \{3, 10\}$ ,  $N(0, 1)$  for  $\beta = 0.5$ , and  $N(0, 2^2)$  for  $\beta = 0$ .

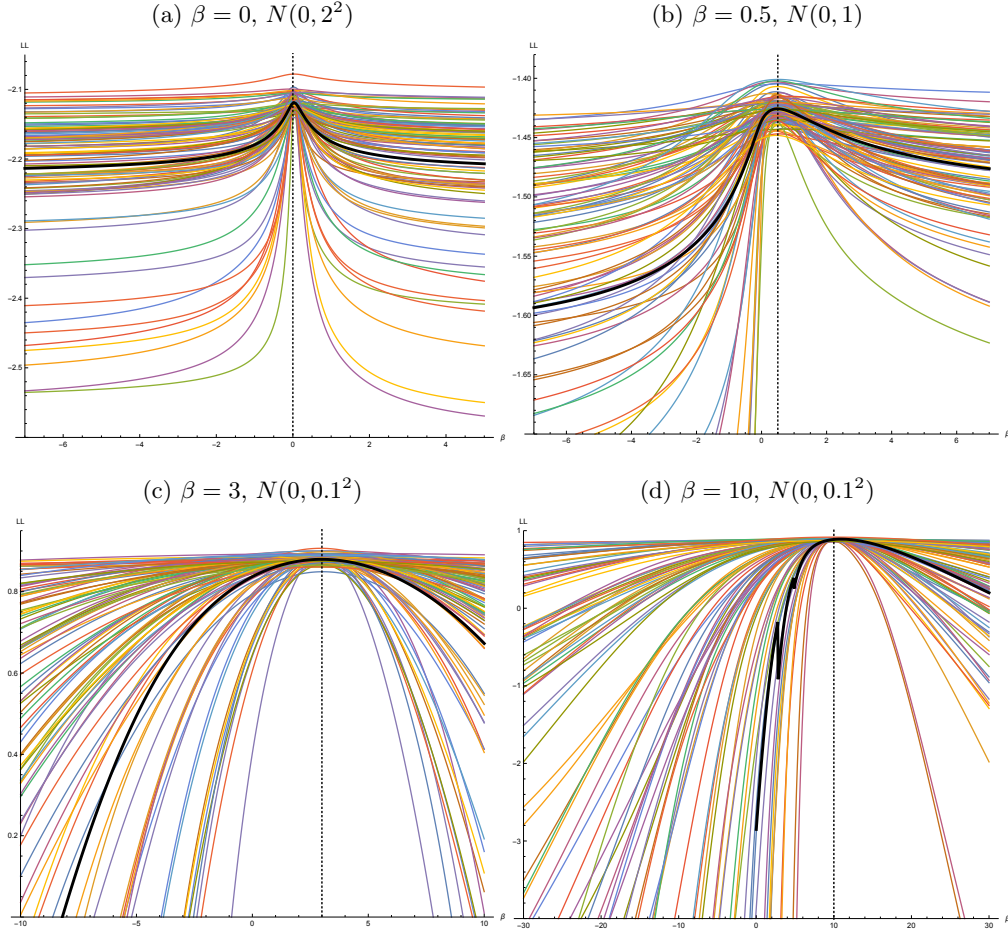
#### 4.3.4. Behaviour of the simulated log-likelihood function

Kristensen and Shin (2012, pg. 80–81) define a set of regularity conditions A.1–A.4 regarding the model and its associated conditional density that ensure sufficiently fast convergence of  $\hat{c} \rightarrow c$  and thus asymptotic equivalence of  $\hat{\theta}$  and  $\tilde{\theta}$ . These conditions basically impose restrictions on the data-generating functions and the conditional density that is being estimated. With regard to data-generating functions, authors argue that the “smoothness conditions are rather weak, and satisfied by most models”. For the conditional density function, they state that “the assumptions are quite weak and are satisfied by many models”. However, for the HAM, we are not able to verify these conditions analytically and we must rely on graphical computational tools. Another important issue regards the identification of parameters in the model assuring uniqueness of the set of estimates.

<sup>17</sup>Moreover, for extreme cases  $\beta$ s =  $\{0, 0.1\}$  noise  $N(0, 2^2)$  seems optimal.



Figure 3: Shape of the simulated log-likelihood function



*Note:* Results based on 100 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distribution. Black dotted vertical lines depict the true  $\beta$ s. Bold black full lines depict sample averages.

For both purposes, we draw the simulated log-likelihood function and verify an existence of a unique maximum. We depict simulated log-likelihood functions for the same four interesting values of the intensity of choice  $\beta \in \{0, 0.5, 3, 10\}$  combined with three specifications of the stochastic noise:  $\epsilon_t \sim N(0, 2^2)$ ,  $\epsilon_t \sim N(0, 1)$  and  $\epsilon_t \sim N(0, 0.1^2)$  as in [Subsection 4.3.2](#). In [Figure 3](#) we clearly observe very smooth shape of the functions over the entire assessed domain with a unique maximum generally shared for all of 100 random runs. Bold black full lines then represent sample averages over these 100 runs and brings the aggregate information,<sup>18</sup> which is, however, obvious also directly from the set of 100 original simulated log-likelihood functions. Based on generally smooth shapes and unique optima of the simulated log-likelihood functions we assume that the

<sup>18</sup>The only violation of the smoothness of the averaged function appears in subpart (d) with relatively high value  $\beta = 10$  where for several runs the model diverges. Disruptions of the averaged function are thus only of the technical origin when depicting, not the feature of the function itself.

regularity conditions are met for the HAM and the identification of parameters is also assured.

#### 4.3.5. Robustness check

To assess the robustness of our general model setting, we contrast the results in Table 10 with several setup modifications. First, in Table 11 in Appendix 1 we consider 10 times higher gross interest rate  $R = 1 + r = 1.001$  representing real market risk free rate. This daily value unrealistically represents circa 28.4% annual risk free interest rate that can nonetheless serve as a useful robustness check. The only considerable effect appears in the increased probability of a ‘NaN’ outcome for the smallest distribution intervals of the stochastic noise  $\epsilon_t$  [see subparts (a), (b), and (c)]. Conversely, for larger distribution intervals of the stochastic noise  $\epsilon_t$  the results are comparable and largely similar. This is another important finding mainly for the empirical application in Section 5 where time series of price deviations are likely to be associated with higher standard deviation of the assumed stochastic market noise. The robustness of the method w.r.t. assumption of the real market risk free rate therefore relaxes the need of a very precise derivation of this parameter for various countries and historical periods and the reasonable approximation  $R = 1 + r = 1.0001$  representing circa 2.5% annual risk free interest rate can be generally used in Section 5.

Second, we test the ability of the estimation method to provide unbiased estimates even if bounds of the parameter space are off-centered, more specifically shifted up by 50% of actual  $\beta$  to  $\langle -0.5\beta, 3.5\beta \rangle$  for  $\beta > 0$  and to  $\langle -0.375, 625 \rangle$  for  $\beta = 0$ . When results of this testing summarized in Table 12 in Appendix 1 are compared to the original results in Table 10, we clearly observe expected shift in the 2.5 (LQ) and 97.5 (HQ) quantiles of the estimate distribution but the ability of the NPSMLE method to reveal true parameter with high precision remains unaffected and the standard deviations are to a great extent similar to the original settings. We therefore verified that there is no need of an excessive precision of the unconstrained pre-estimation via which we define bounds of the parameter space for the constrained optimisation.

We further test how is the NPSMLE method performance affected by assumption of another than normal distribution of the stochastic noise  $\epsilon_t$ . For that purpose we select uniform distribution both for its simplicity as well as for its feature of being the maximum entropy probability distribution of its family of symmetric probability distributions. As the assumption of normal distribution of stock market noise seems reasonably realistic (see discussion in Subsection 4.1), the uniform distribution provides a less realistic candidate for the robustness check. We consider two principles for comparison of our original results with the results assuming the uniform distribution—the first based on almost identical covered intervals (Table 13 in Appendix 1) and the second based on the same variance (Table 14 in Appendix 1). We therefore define a grid of 10 uniform noise specifications

1. that cover the same intervals as are covered by their respective normal counterparts by the 99.74% of the probability mass;<sup>19</sup>
2. with the same variances as their respective normal counterparts.<sup>20</sup>

Contrasting original results in Table 10 with results based on the uniform distribution of the stochastic noise, we basically verify the assumption of Kristensen and Shin (2012) that  $\mathcal{F}_\epsilon$  can be

<sup>19</sup>I.e.  $\langle -3SD, 3SD \rangle$ , where  $SD$  stands for the standard deviation of the respective normal noise specification.

<sup>20</sup>Variance (or the second centralized moment) of the continuous uniform distribution is defined as  $\frac{(b-a)^2}{12}$ , where  $a$ ,  $b$  are the minimum and maximum values of the distribution’s domain.

Table 4: Results for  $\beta$  estimation with normal noise, fixed  $g_h$  &  $b_h$ 

$\beta$	(a) $\hat{\beta}, N(0, 0.1^2)$						(b) $\hat{\beta}, N(0, 1)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	.00	.13	-.26	.28	0%	-.00	-.00	.05	-.10	.09	0%
.1	.11	.10	.12	-.10	.30	0%	.10	.10	.06	-.02	.22	0%
.5	.49	.49	.11	.26	.71	0%	.50	.51	.17	.19	.85	0%
1	1.00	1.00	.07	.86	1.14	0%	1.00	1.01	.18	.67	1.40	0%
3	3.01	3.00	.13	2.75	3.23	0%	3.03	3.12	.57	2.49	5.28	0%
5	5.00	5.00	.23	4.56	5.47	0%	5.19	6.42	2.70	4.25	13.65	4%
10	10.00	10.00	.12	9.77	10.23	0%	11.04	12.28	4.64	6.07	24.11	19%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distributions of given parameters,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

any known distribution. The overall result are largely similar and the observed differences can be attributed mainly to different shapes of the normal and uniform distribution. In the case of identical covered intervals (Table 13) we observe slightly lower probability of ‘NaN’ outcome occurrence for smallest intervals [subparts (a), (b), and (c)], but higher for largest intervals [subparts (i) and (j)]. In the case of identical distribution (Table 14), the ‘NaN’ outcome occurrence is comparable with the normal distribution. In both cases the efficiency of estimates tends to be higher than using the normal distribution and intervals between the 2.5 (LQ) and 97.5 (HQ) quantiles are narrower in majority of specifications. We can assign this to differences in shape of compared distribution—the highest probability density of the normal distribution around the zero mean combined with possibility of extreme observations, both apparently negatively affecting the efficiency of estimates.

One of the most challenging concepts in the estimation method setting is the repeated random generation of parameters  $g_h$  and  $b_h$  for each from 1000 runs. In Table 4 we abandon this setup feature, fix parameters  $g_h$  and  $b_h$  randomly before the loop and use the very same figures for all 1000 runs. Although the repeated random generation of trend and bias parameters for each run is one of the robustness cornerstones of the analysis in Section 4, it is of our interest to provide the comprehensive picture of the NPSMLE method performance. Next in this section, we only compute and depict results for two specification of the stochastic noise which appeared the most useful based on results in Table 10, namely  $N(0, 0.1^2)$  and  $N(0, 1)$ . Comparing results in Table 4 with respective counterparts Table 10 [subparts (h) and (i)], aside a minor reduction of the ‘NaN’ emergence probability we observe overall significant reduction of standard deviations of  $\beta$  estimates. Fixing trend a bias parameters thus naturally makes the system more predictable and leads to more efficient estimates.

Tables 15, 16 in Appendix 1 report results of another robustness check of the methodology focused on various distributions of belief parameters  $g_h$  and  $b_h$  (see Equation 18). To recap, in the general model we follow the previous work of Barunik et al. (2009); Vacha et al. (2012); Kukacka and Barunik (2013) and thus trend parameters  $g_h$  are drawn from the normal distribution  $N(0, 0.4^2)$  and bias parameters  $b_h$  are drawn from the normal distribution  $N(0, 0.3^2)$ . Here we relax this assumption using a range of reasonable variances defining the distribution of beliefs’ parameters:  $\{0.1^2, 0.2^2, 0.3^2, 0.4^2, 0.6^2, 0.8^2, 1, 1.2^2\}$ .

Analysing results in Tables 15 and 16, we observe the general ability of the method to reveal accurately the true value of the intensity of choice  $\beta$  for the vast majority of combinations of the

simulation grid as well as for both noise specifications. On the other hand, in specific cases we can observe signs of an upward bias, values on the border of statistical significance, or considerable probability of ‘NaN’ outcome emergence. The upward bias is observable only in [Table 16](#) associated with the larger specification of the stochastic noise  $N(0, 1)$  and for higher values of  $\beta$ . However, the upward bias tendency disappears with increasing variances of the distribution of beliefs’ parameters. This is another example of somehow puzzling behaviour for which we cannot find any obvious explanation. The problem with statistical insignificance of estimates is naturally mostly evident for distribution specification with small variances and small values of  $\beta$ , both technically inhibiting the dynamics of the model. Increasing the variance of beliefs’ distribution associated with higher values of randomly generated belief parameters  $g_h$  and  $b_h$ , we generally obtain a richer model dynamics which can be more simply and more efficiently estimated. However, there is a crucial trade-off in form of model divergence and results censorship by ‘NaN’ outcomes reported in the NN column. This situation is generally more likely for higher values of the intensity of choice  $\beta$  as well as for the higher potential values of belief parameters  $g_h$  and  $b_h$ . For instance, the two highest values of our simulation range:  $\{1, 1.2^2\}$  induce serious results censorship (up to 96%) for the majority of values from the discrete range of  $\beta$ s, particularly when combined the stochastic noise  $N(0, 1)$  that is generally associated with considerably higher probability of ‘NaN’ outcome emergence. Small stochastic noise specifications generally do not exhibit signs of upward biases but can be also largely affected by ‘NaN’ outcomes. We refer the reader to additional results of the stochastic noise specification  $N(0, 10^{-12})$  and  $N(0, 10^{-14})$  summarized in [Table 17](#), [Table 18](#) in [Appendix 1](#). The inference based on such filtered result is invalid and we concur the discussion about the ‘NaN’ outcomes originated in [Subsection 4.3.3](#) via describing this third<sup>21</sup> aspect triggering the divergence of the model leading to ‘NaN’ outcome of the NPSMLE procedure. We observe a somewhat nontrivial complex interplay between the intensity of the stochastic noise, estimation efficiency, and probability of ‘NaN’ outcomes. Higher probabilities are associated both with tiny noises  $N(0, 10^{-14})$  and  $N(0, 10^{-12})$  (see [Tables 17](#) and [18](#)) as well as with larger noise  $N(0, 1)$  ([Table 16](#)) when compared to  $N(0, 0.1^2)$  ([Table 15](#)). However, the mitigating impact on the results censorship in specific setups is far from solving the censorship issue. When comparing results of  $N(0, 0.1^2)$  and  $N(0, 1)$  in terms of efficiency, one would conclude that the wider distribution interval  $\epsilon_t \sim N(0, 1)$  increases the efficiency of estimates. But the opposite holds when comparing  $N(0, 10^{-14})$  and  $N(0, 10^{-12})$  in [Appendix 1](#). We observe this partly ambiguous relation between stochastic noise distributions intervals and values of  $\beta$  also in the original results in [Table 10](#).

To sum up hitherto findings regarding the robustness of the NPSMLE method w.r.t. various setup specifications, we face an interesting ‘two-sided’ trade-off. Basically, we are able to estimate relatively well a model exhibiting reasonably rich dynamics. This is, however, on the one hand inhibited assuming:

1. low values of the intensity of choice  $\beta$ ,
2. small distribution intervals of the stochastic noise  $\epsilon_t$ ,
3. or distribution specifications of belief parameters  $g_h$  and  $b_h$  with small variances,

producing insufficient dynamics or fragile stability of the system. But the other hand also by:

1. too high values of the intensity of choice  $\beta$

---

<sup>21</sup>Along with very small distribution intervals of the stochastic noise  $\epsilon_t$  and high values of the intensity of choice  $\beta$ .

Table 5: Results for  $\beta$  estimation with various combined noises I.

$\beta$	(a) $\hat{\beta}$ , $\epsilon_t \sim N(0, 0.1^2)$ , $\{\varepsilon_i\}_{i=1}^N \sim N(0, 1)$						(b) $\hat{\beta}$ , $\epsilon_t \sim U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1})$ , $\{\varepsilon_i\}_{i=1}^N \sim U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.29	-.47	.47	0%	-.01	-.00	.29	-.48	.47	0%
.1	.10	.10	.11	-.09	.29	0%	.10	.10	.11	-.09	.29	0%
.5	.50	.50	.58	-.47	1.45	0%	.51	.51	.58	-.47	1.46	0%
1	.97	.99	1.13	-.86	2.89	0%	.99	.98	1.12	-.90	2.87	0%
3	2.87	2.84	3.30	-2.67	8.60	0%	2.87	2.86	3.43	-2.77	8.72	0%
5	4.89	4.87	5.31	-4.48	14.35	0%	5.24	5.32	5.70	-4.69	14.63	0%
10	10.96	11.42	10.56	-8.40	28.99	0%	11.19	10.21	11.65	-9.95	29.17	1%
(c) $\hat{\beta}$ , $\epsilon_t \sim N(0, 1)$ , $\{\varepsilon_i\}_{i=1}^N \sim N(0, 0.1^2)$							(d) $\hat{\beta}$ , $\epsilon_t \sim U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2})$ , $\{\varepsilon_i\}_{i=1}^N \sim U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1})$					
100%							100%					

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from the same distributions with different variances,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

2. large distribution intervals of the stochastic noise  $\epsilon_t$ ,
3. or distribution specifications of belief parameters  $g_h$  and  $b_h$  with large variances,

potentially causing model to diverge and therefore censoring the results. This seems to be another of challenging issues for the empirical application of the NPSMLE method.

The last part of [Subsection 4.3.5](#) addresses an important question of what happens if wrong stochastic noise assumption is used to perform the NPSMLE? This can either be a right distribution, but with wrong parameters, or a completely different distribution. This question is especially important w.r.t. empirical application in [Section 5](#) because in real world data we are rarely able to ascertain proper noise. For the purpose of analysing this issue, we present results of various combinations of different distributions used for random generation of stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$ . In [Table 5](#) we report the case where stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  are drawn from the same distributions with different variances, [Table 19](#) in [Appendix 1](#) then displays results when different distributions with various variances (same as well as different) are combined together. Basically, we use combinations of normal and uniform distributions and for different variances we use specifications with 10 times higher or lower values. Conclusions for this robustness check are very clear and can be summarized into several points:

1. when a distribution with (10 times) higher variance is used for generating stochastic noise  $\epsilon_t$  then for kernel approximation of the conditional density  $c_t(x|v; \theta)$  via  $\{\varepsilon_i\}_{i=1}^N$ , the NPSMLE method is inapplicable as this situation leads to complete ‘NaN’ outcome. The reason is very different from a usual ‘NaN’ occurrence caused by HAM divergence. In this case, apparently, the method itself is not able to approximate the true conditional density using  $\{\varepsilon_i\}_{i=1}^N$  generated from 10 times ‘smaller’ distribution. This holds irrespective of whether same or different distributions are used [see subparts (c) and (d) of [Table 5](#) and (g) and (h) of [Table 19](#)];
2. when a distribution with (10 times) lower variance is used for generating stochastic noise  $\epsilon_t$  then for kernel approximation of the conditional density  $c_t(x|v; \theta)$  via  $\{\varepsilon_i\}_{i=1}^N$ , the NPSMLE



method works from the technical point of view but produces completely statistically insignificant estimates and very random uniform distribution of estimated values as captured in HQ and LQ columns where the quantile values almost copy the borders of the parameter space, this also holds irrespective of whether same or different distributions are used;

3. when different distributions with the same variances are used [compare subparts (a), (b) and (c), (d) of Table 19], the order and shapes of distributions matter considerably. In our case, if stochastic noise  $\epsilon_t$  is generated from the normal distribution and  $\{\epsilon_i\}_{i=1}^N$  is generated from the uniform distribution [subparts (c) and (d)], we obtain considerably better estimates with markedly lower standard deviations than if the order is opposite [subparts (a) and (b)]. As both distributions have identical mean and variance, there is not much than shape of the distribution defined by higher moments to make difference.

These findings very strongly confirm the need of a proper noise specification for the empirical application of NPSMLE which is at the same moment one of the most important findings from the analysis of the original results from Table 10.

#### 4.3.6. 2-type model estimation

An important advantage of FABMs is that their dynamics is mostly driven by a few crucial parameters. As a result, we might promisingly attempt to estimate all essential coefficients simultaneously and thus we do not need any rigorous criteria for selection. In the Brock and Hommes (1998) setting we select estimated parameters consistently with the current literature (see Tables 1 and 2), i.e. the key switching parameter  $\beta$  and the behavioural belief coefficients. The other coefficients, e.g. the risk aversion  $a$ , the conditional variance of excess returns  $\sigma^2$ , or the risk free rate  $R$  are simplified already in the original model as constants and shared by all investor types. The model is then theoretically derived based on those assumptions. These parameters only influence the absolute values of the profitability measures  $U_h$  but not their relative proportions ( $R$  additionally a little bit adjusts the model output  $x_t$ ). Thus we can naturally consider them not influencing dynamics of the model as described in Subsection 4.1.

A natural subsequent step of the NPSMLE method testing is thus a multiple parameter estimation in which we simultaneously estimate the intensity of choice  $\beta$  and agents' belief coefficients  $g_h$  and  $b_h$  defining individual trading strategies in the 2-type and the 3-type models for which both theoretical as well as empirical rationale exists in the current literature as indicated in Chen et al. (2012, pg. 191, 207). With reference to Biondi et al. (2012, pg. 5534), "it has been advocated that the two broad categories of chartism and fundamentalism account for most of possible investment strategies". The aim of this analysis is to assess the performance of the NPSMLE method in estimating other model parameters then solely the intensity of choice  $\beta$ .

First, we study the most simple system consisting of two trading strategies, where fundamental strategy again appears in the market by default ( $g_1 = b_1 = 0$ ). Based of the knowledge gained in Subsection 4.3.1, we define a discrete grid of combinations of the true intensity of choice  $\beta$  and the chartistic beliefs  $g_2$  and  $b_2$  representing the second-type trading strategy to cover a purposeful range of values w.r.t. issues studied in the previous sections. To keep a reasonable number of combinations and lucidity of results, we opt for  $\beta = \{0, 0.5, 3, 10\}$ . In defining a grid of chartistic beliefs, we also cover various combinations of trend following ( $g_2 > 0$ ), contrarian ( $g_2 < 0$ ), upward-biased ( $b_2 > 0$ ), and downward-biased ( $b_2 < 0$ ) strategies based on multiples of standard deviations from the general model setting:  $0.5\times, 1\times, 2\times, 3\times$ . We refer the reader to

the first column of [Table 20](#) in [Appendix 1](#) for detailed specification. We again employ only two specification of noise, namely  $\epsilon_t \sim N(0, 0.1^2)$  and  $\epsilon_t \sim N(0, 1)$ .

#### 4.3.7. Quantitative results

In [Table 20](#), [Table 21](#) in [Appendix 1](#) we summarise the simulation results. Basically, we are able to confirm all main findings from the single parameter  $\beta$  estimation simulation analysis. First, the method is generally able to reveal accurately the true values (see columns reporting sample medians and means denoted ‘Md’ and ‘Mn’, respectively) of estimated parameters also in the 3-parameter simultaneous estimation case. Especially belief coefficients  $g_2$  and  $b_2$  [subparts (b) and (c)], that are of central importance in this section, are estimated overall significant and with almost surprisingly high precision. The estimation precision of the  $\beta$  parameter is not directly comparable to previous results as the setting of the 2-type model is different from the general model. Notwithstanding, we still get generally conformable figures. Second, the ‘two-sided’ trade-off (see [Subsection 4.3.5](#)) related to a reasonably rich model dynamics restricted from both sides plays an important role. We again clearly observe the relative estimation inefficiency in case of setting combinations with small values of the intensity of choice  $\beta = 0, 0.5$ , especially when combined with small values of belief coefficients  $g_2$  and  $b_2$  (the upper half of [Table 20](#) and [Table 21](#)). On the other hand, for high values of belief coefficients  $g_2$  and  $b_2$  (the bottom half of [Table 20](#) and [Table 21](#)) we experience considerable number of model overflows leading to significant censorship of results and thus generally confirming findings from  $\beta$  estimation exercise w.r.t. various distributions of beliefs parameters ([Tables 15](#) and [16](#)). These effects to a large extent prohibit serious results interpretation for reported combination with the smallest values of belief coefficients  $g_2 = \pm 0.2$  and  $b_2 = \pm 0.15$  as well as for the combination with the highest values of belief coefficients  $g_2 = \pm 1.2$  and  $b_2 = \pm 0.9$ . We also observe the nontrivial complex interplay between the intensity of the stochastic noise, estimation efficiency, and probability of ‘NaN’ outcomes [in subparts (d), ‘NN’ column] detected within the analysis of the general model. Here, a stabilising effect is associated with the wider noise interval  $N(0, 1)$  leading to a significant decrease of the number of model overflows (on the contrary, in the case of 3-type model estimation in [Subsection 4.3.10](#), the stability is higher for the narrower noise  $N(0, 0.1^2)$  specification). This may paradoxically lead to a seemingly lower efficiency as the divergent runs are not filtered out—this effect is e.g. observable in [Table 21](#) for combinations with  $\beta = 10$ ,  $g_2 = \pm 0.4$ , and  $b_2 = \pm 0.3$  where  $\beta$  estimates are often biased upwards. The standard deviations of these empirical estimates is significantly larger compared to [Table 20](#) due to the effect of estimates close to the upper bound of the parameter space that would otherwise be likely filtered out in the case of smaller stochastic noise in the system. We further observe a prevailing upward bias tendency in  $\beta$  estimates for smallest values of belief coefficients  $g_2$  and  $b_2$ . When it comes to efficiency of  $g_2$  and  $b_2$  estimates, the setup with noise interval  $N(0, 0.1^2)$  produces markedly more precise estimates but the overall statistical significance of estimates is apparent for both setups.

Although we report all combination of trend following ( $g_2 > 0$ ), contrarian ( $g_2 < 0$ ), upward-biased ( $b_2 > 0$ ), and downward-biased ( $b_2 < 0$ ) strategy specification mainly from technical reasons and the analysis of the model dynamics goes beyond the scope of this paper, we can observe some patterns regarding the beliefs’ combination. For instance the efficiency of the estimation is considerably higher for trend following ( $g_2 > 0$ ) beliefs than for contrarians ( $g_2 < 0$ ), but setups with trend following beliefs are more vulnerable to overflows causing ‘NaN’ outcomes. On the other hand the direction of the bias parameter  $b_2$  does not seem to play any important role based on reported results.

At the end of the 2-type model analysis, it is, however, important to stress that this is the

extreme case of the most simple setting more vulnerable to potential extreme model dynamics. Nonetheless, this simple setting is the cornerstone in the current HAM estimation literature (see [Section 2](#)) and therefore it needs to be properly elaborated also for the new NPSMLE method. The favourable results presented above give promise for the function of the method also in more complex settings.

#### 4.3.8. Behaviour of the simulated log-likelihood function

To verify smoothness conditions and identification of parameters in the 2-type model estimation case, we aim at depicting shape of simulated log-likelihood functions also for the simultaneous estimation of three parameters. As we can hardly demonstrate the 4D shape of the resulting simulated log-likelihood function, we depict sub-log-likelihood functions in 2D and 3D making out the global visualisation when combined together.

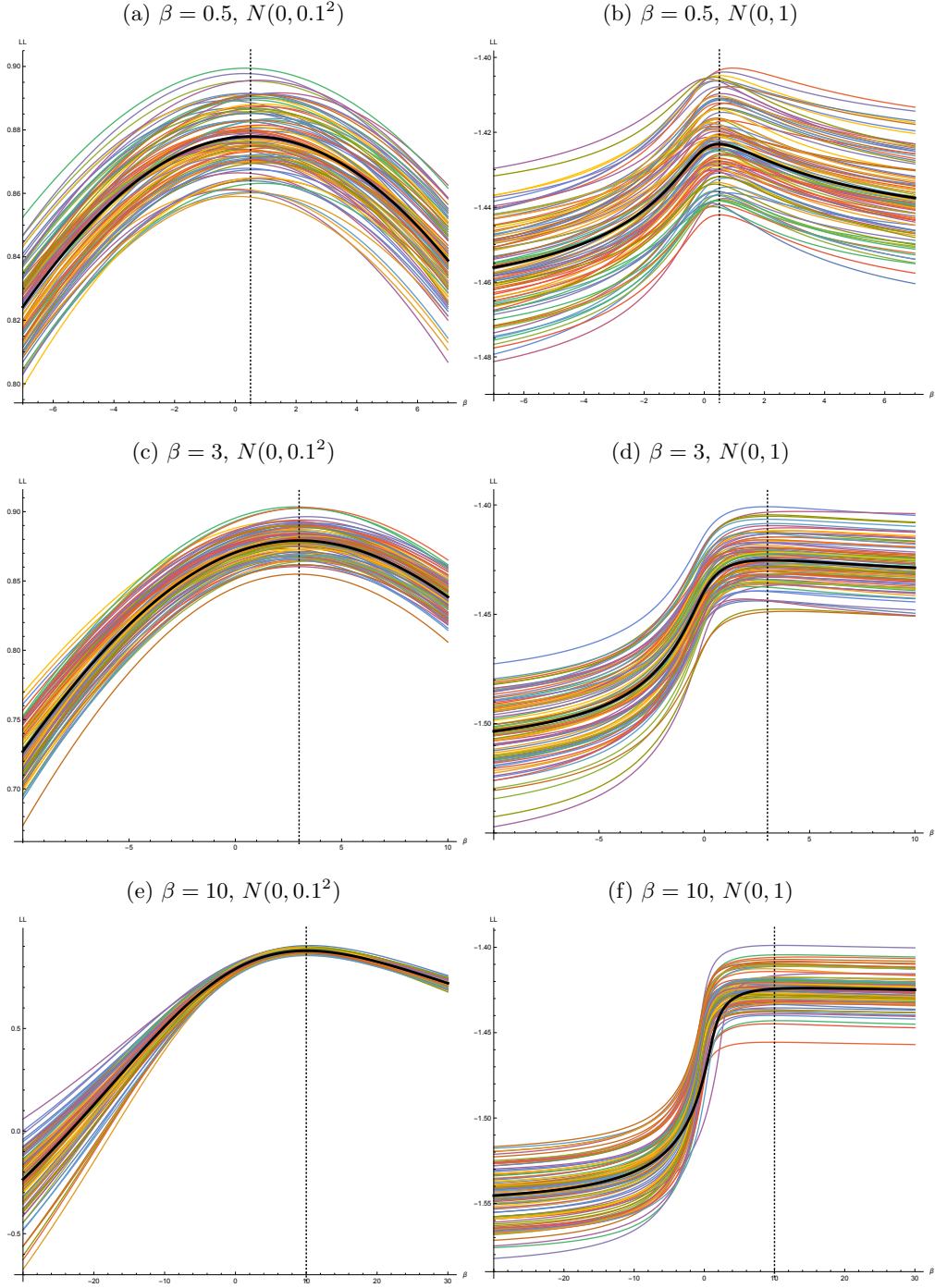
Figures 4 and 10 in [Appendix 2](#) demonstrate simulated 2D sub-log-likelihood functions for the single parameters intensity of choice  $\beta \in \{0.5, 3, 10\}$ , trend coefficient  $g_2 = 0.4$ , and bias coefficient  $b_2 = 0.3$  estimation (keeping the two others fixed) combined with the two most ‘successful’ specifications of the stochastic noise  $\epsilon_t \sim N(0, 0.1^2)$  and  $\epsilon_t \sim N(0, 1)$  found in [Subsection 4.3.3](#). We again simply observe very smooth shapes and unique maxima generally shared for all random runs. Moreover, consistent to [Subsection 4.3.3](#) findings, small  $\beta = 0.5$  is more precisely detectable assuming stochastic noise  $\epsilon_t \sim N(0, 1)$  [subfigure (b) of [Figure 4](#)], higher  $\beta \in \{3, 10\}$  assuming stochastic noise  $\epsilon_t \sim N(0, 0.1^2)$  [subfigures (c) and (e) of [Figure 4](#)], upward bias tendency of  $\hat{\beta}$  is clear for the stochastic noise  $\epsilon_t \sim N(0, 1)$  in subfigures (d) and (f) of [Figure 4](#) due to very flat shape of the log-likelihood function above the positive subpart of the domain, and finally the belief parameters  $b_2 = 0.3$  and  $g_2 = 0.4$  are very well detected in both cases, however the performance of estimators is more efficient for smaller stochastic noise  $\epsilon_t \sim N(0, 0.1^2)$ . Next, in [Figure 11](#) in [Appendix 2](#) we visualise 3D simulated sub-log-likelihood functions based on all possible combinations of three parameters of interest, keeping one of them fixed, for the model setting with:  $\beta = 0.5$ ,  $g_2 = 0.4$ ,  $b_2 = 0.3$  and both stochastic noise specifications. The smoothness of the surface generally keeps retained also in the 3D visualisations and regions of possible maxima are well detectable via red color although the 3D depiction cannot provide such detailed and ‘zoomed’ view as the 2D visualisations in [Figures 4](#) and [10](#). For some combinations of parameters the model is numerically unstable and thus for specific subsets of the domain plane the shape is not well defined. However, these areas are always far from maxima regions. Parameters  $g_2$  and  $b_2$  seem to be relatively well identified which is further confirmed quantitatively in [Subsection 4.3.7](#). As expected, the most challenging is revealing the  $\beta$  coefficient in which direction the surface is very flat for a large interval of the domain. These findings are largely in accord with conclusions of [Bolt et al. \(2014, pg. 15\)](#) and [Hommes and Veld \(2015\)](#) who claim that “the other parameters can to a large extent compensate for changes in  $\beta$ ” and report very flat shape of the likelihood function for the intensity of choice selection. In any case, based on these results we again generally assume that the regularity conditions are met and the identification of parameters is assured also for the 2-type model estimation.

#### 4.3.9. Likelihood-ratio test

Previous sections have shown that the NPSML estimation method does a fairly good job in distinguish between various  $\beta$ s. As a next step we might be interested how capable the estimation method is in a rigorous statistical comparison between static and switching models. As the static version of the model with  $\beta = 0$  and the switching version are nested, i.e. the less complex static



Figure 4: Simulated sub-log-likelihood functions for  $\beta$  estimation



*Note:* Results based on 100 random runs,  $g_2 = 0.4$ ,  $b_2 = 0.3$ , number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distribution. Black dotted vertical lines depict the true  $\beta$ s. Bold black full lines depict sample averages.

model is derived via a restriction on  $\beta$  from the switching model, we can apply the usual likelihood ratio test to assess the relative goodness of fit between models with and without switching. For this purpose [Table 20](#) and [Table 21](#) further display information about tests of model fit in subpart (d). ‘L-rat’ column denotes the likelihood ratio of the null static (i.e. restricted) model vs. the alternative switching model, ‘ $2\Delta LL$ ’ is the test statistics of the log-likelihood ratio test being approximately  $\chi^2$  distributed with 1 degree of freedom (because only the switching parameter is restricted), and ‘p-v’ is related p-value of the test. Application of the likelihood ratio test seems natural in this situation, nevertheless the Monte Carlo simulation framework brings several imperfections. On the one hand, simulated data smartly avoid the problem of model misspecification, on the other hand the goodness of fit test is designed rather for comparison based on a single empirical dataset. In our situation, new variables and a different dataset are randomly generated in each of 1000 random runs and the test is then based on the aggregate mean values over all runs. But since the standard deviation of individual maximum log-likelihoods is negligible compared to their value (around 1%), aggregation does not cause any appreciable loss of information. A second imperfection is related to the relative flatness of the log-likelihood function in the dimension of the restricted switching coefficient  $\beta$  for a large interval of the domain (see detailed discussion in [Subsection 4.3.8](#)). Although the estimation method detects precisely the true intensity of choice especially for combinations of higher values of  $\beta$ , stronger strategies, and lower stochastic noise specification, due to flat likelihood the test exhibits only a moderate capability of distinguishing between the restricted and the unrestricted model. Translated into p-values of the test to reject the null of the static model, in the most distinct cases the value reaches 30% which is far above usual econometric levels. Inspecting subpart (d) of [Table 20](#) and [Table 21](#), we observe expected behaviour but generally low power of the test. For all true  $\beta = 0$ , the likelihood ratio is equal to 1 and the p-value remains 100%. Increasing true  $\beta$  and strength of strategies, the likelihood ratio and the p-value naturally decrease (as the test statistics ‘ $2\Delta LL$ ’ increases), but the pace of the progress is low for the selected range of  $\beta$ s.

#### 4.3.10. 3-type model estimation

Results of simultaneous estimation of 5 parameters in the 3-type model including three basic strategies: fundamental represented by  $g_1 = b_1 = 0$  and two chartistic defined in the [Table 22](#), [Table 23](#), can be found in [Appendix 1](#). We keep the strategy of defining a grid of chartistic beliefs from [Subsection 4.3.6](#) based on various combinations of trend following ( $g_2 > 0$ ), contrarian ( $g_2 < 0$ ), upward-biased ( $b_2 > 0$ ), and downward-biased ( $b_2 < 0$ ) strategies and the same multiples of standard deviations from the general model setting. Conclusion are generally in accord with the results of the 2-type model estimation, the difference is mainly in efficiency of estimates that is by nature lower than for the 2-type model. However, combining two chartistic strategies we also gain some new knowledge about the system behaviour. For instance, in case of a combination of two trend following strategies ( $g_2 > 0, g_3 > 0$ ), it is rather complicated for the NPSMLE method to distinguish between impacts of these two strategies leading to lower statistical significance of both estimates compared to other combination of trend following and contrarian strategies. Conclusions regarding the nontrivial interplay between the intensity of the stochastic noise, estimation efficiency, and the probability of ‘NaN’ outcomes seem to be somewhat mixed in this markedly more complex case as the estimation efficiency is comparable for both stochastic noise specifications  $N(0, 0.1^2)$  and  $N(0, 1)$ , and the stability of the system in terms of ‘NaN’ outcomes is higher for  $N(0, 1)$  which is the opposite tendency than observed within the analysis of the 2-type model estimation.

## 5. HAM estimation on empirical data

Equipped with the knowledge from the Monte Carlo study in [Section 4](#), we broaden the topic via an empirical application and estimate the [Brock and Hommes \(1998\)](#) model using cross section of world stock markets. We analyse S&P500 and NASDAQ for the U.S., DAX and FTSE and for Europe, NIKKEI 225 and HSI for Japan and Hong Kong, respectively.

### 5.1. The estimation setting

Compared to the simulation study in [Section 4](#), the setting of models estimated using real data is less challenging and in terms of statistical validity also less computationally demanding. On the contrary, the estimation algorithm is a bit technically more complicated as the structure of the real world data is far away from the regularity of the simulated dataset. Concurring findings about the sufficient setting of the NPSMLE method from [Section 4](#), we compute results for 1000 random runs, number of observations  $t = 5000$ , and the kernel approximation precision is set to  $N = 500$ . On the other hand, because of a problematic numerical stability of the model when real data analysis is introduced we increase the number of starting points for the numerical optimisation to 8.<sup>22</sup> The other setting remains the same as defined in [Subsection 4.1](#). More importantly, as the effect of the more complicated structure of the code with multiple initial points, the parallelisation of the computational procedure cannot be maintained via the *MATLAB* because of technical assumptions of the parallelisation function. Therefore we are left with simple and relatively slow standard one-core type of calculations.

Moreover, based on main findings of [Section 4](#), the proper intensity of stochastic market noise is crucial for the correct function of the NPSMLE method. A wrong stochastic noise specification is likely to influence the behaviour of the system and validity of results to a great extent. Hence we do not longer use the grid strategy to ensure the robustness of result as e.g. the stochastic noise intensity on various real markets is immensely unpredictable variable. Leaving this idea, we instead add the intensity of the stochastic market noise to the list of estimated parameters. Generally, we thus apply a simultaneous unconstrained multivariable function estimation of all interesting parameters: agents' belief coefficients defining individual trading strategies  $g_h$  and  $b_h$ , the intensity of choice  $\beta$ , and the intensity of market noise, which is defined as a fraction of the standard deviation of the noise term and the standard deviation of the data and denoted as *noise intensity*.

First, we estimate the most simple 2-type model including two basic strategies only—fundamental one represented by implicitly defined  $g_1 = b_1 = 0$  and chartistic one which is to be estimated. Within this setting, we simultaneously estimate four parameters of interest— $\beta$ ,  $g_2$ ,  $b_2$ , and the *noise intensity*. To support the numerical stability of the estimated system, we constrain the intervals for the starting points random generation to  $\langle -0.5, 0.5 \rangle$  for  $\beta$ ,  $\langle 1.3, 2.3 \rangle$  for  $g_2$ ,  $\langle -0.2, 0.2 \rangle$  for  $b_2$ , and  $\langle 0.4, 0.9 \rangle$  for the *noise intensity*. Nonetheless, as the algorithm is designed to find an optimum of an unconstrained multivariable function, it can freely leave these initial intervals during the search procedure.

We then continue with the estimation of the 3-type model including three basic strategies—fundamental and two different chartistic strategies which are to be estimated. Based on results of the 2-type model estimation, we assume zero bias of both the trend following as well as the

---

<sup>22</sup>Based on a preliminary analysis, 8 starting points are sufficient to filter out issues related to numerical instability on the system.

contrarian strategy, i.e.  $b_2 = b_3 = 0$ . The simultaneous estimation of four parameters of interest— $\beta$ ,  $g_2$ ,  $g_3$ , and the *noise intensity*—technically requires a modification of the algorithm setting to the constrained multivariable function estimation as the two different strategies—the trend following  $g_2 > 0$  and the contrarian  $g_3 < 0$  need to be strictly distinguished using the following constraints:  $g_2 \in \langle 1.8, 2.8 \rangle$ ,  $g_3 \in \langle 0, -0.5 \rangle$ .

## 5.2. Fundamental price approximation

Approximation of the fundamental price is inevitably the most ‘challenging’ issue of the entire empirical section. Unfortunately, in the original framework of the [Brock and Hommes \(1998\)](#) asset pricing model of a simple stylised stock market there is no hint about how the empirical fundamental value might be derived.<sup>23</sup> Thus, we are left with the existing literature and following [Winker et al. \(2007\)](#); [ter Ellen and Zwinkels \(2010\)](#); [Huisman et al. \(2010\)](#), the fundamental price is approximated as a Moving Average (MA) value. [Winker et al. \(2007\)](#) assume as the proxy for the fundamental price a MA over the last 200 observations of the DM/USD exchange rate time series for the period 1991/11/11 to 2000/11/9. [ter Ellen and Zwinkels \(2010\)](#) use the MA of the Brent and WTI Cushing oil monthly USD prices over 24 months, i.e. from 1984/1 to 2009/8. [Huisman et al. \(2010\)](#) employ the MA of European forward electricity daily historical prices over 3 year for the base-load calendar year 2008 forward contracts. Authors set the MA window to 3 as a calibration result of the optimal length.

Long-term and short-term MAs are also commonly used by practitioners in trading to extrapolate divergence from the fundamental value in technical analysis. Since the fundamental value of stocks is essentially unknown, market practitioners often tend to at least estimate whether the stock is over or under-valued, whether the possible mispricing is small or large, and whether the gap is going to increase or whether a soon correction is more likely. As the [Brock and Hommes \(1998\)](#) model is also formulated in deviations from the fundamental price, the MA approach seems to be one of reasonable guidances. The MA filtering is the cornerstone of technical analysis and therefore widely used by active traders: [Allen and Taylor \(1990, pg. 50\)](#) present empirical evidence on the perceived importance of technical analysis among London foreign exchange dealers and refer to prevalent mechanical indicators such as trend-following rules: “buy when a shorter MA cuts a longer MA from below”. [Taylor and Allen \(1992\)](#) survey chief foreign exchange dealers operating in London and report that 64.3% of organisations use MAs and/or other trend-following analytical techniques. [Brock et al. \(1992, pg. 1735\)](#) refer to MA technical rules as to one of the two simplest and most widely used: “when the short-period MA penetrates the long-period MA, a trend is considered to be initiated”. [Lui and Mole \(1998, pg. 541, 535\)](#) repeat largely analogical survey as [Taylor and Allen \(1992\)](#) among Hong Kong foreign exchange dealers and report that MAs “are seen to be the most useful technical technique at all three horizons” (i.e. intraday, intramonth, > 1 month) and that “technical analysis is considered ... significantly more useful in predicting turning points”. [Goldbaum \(1999, pg. 70, 71\)](#) describes the way how in practice the MA trading

---

<sup>23</sup>In contrast, another class of HAMS of FOREX markets successfully utilises the Purchasing Power Parity between two countries as the approximation of the fundamental value of the currency exchange rate [see e.g. [Vigfusson \(1997\)](#); [Westerhoff and Reitz \(2003\)](#); [Manzan and Westerhoff \(2007\)](#); [Wan and Kao \(2009\)](#); [Goldbaum and Zwinkels \(2014\)](#); [Verschoor and Zwinkels \(2013\)](#)]. [Boswijk et al. \(2007\)](#) and [de Jong et al. \(2009a\)](#) employ the static Gordon growth model for equity valuation proposed by [Gordon \(1962\)](#), which is, however, infeasible for the empirical validation of the original [Brock and Hommes \(1998\)](#) model. Some other papers simply use a RW formula to drive the fundamental price ([De Grauwe and Grimaldi, 2005, 2006](#); [Winker et al., 2007](#); [Franke, 2009](#)).

rules translate into buy-sell indicators: “when the short period moving average, say the average price of the security over the last five trading days, rises above the long period moving average, say the average of the price over the last 200 trading days, this is a buy indicator. When the short period moving average drops below the long period moving average, this is a sell indicator”. Sullivan et al. (1999, 1999) summarise that “MA cross-over rules . . . are among the most popular and common trading rules discussed in the technical analysis literature”. To quote from Isakov and Hollistein (1999), “the most popular moving average rule used is (1,200), where the short period is one day (in fact it is the index itself) and the long period is 200 days (almost a year)”. According to authors, “the academic literature has shown that the best results were obtained when the short average is one day”. Closely related to our work, “motivated by the popularity of MA strategies in real markets and empirical studies” Chiarella et al. (2006, pg. 1748) propose a model in which the demand of chartists is determined by the difference between a long-term MA and current market price.

For the MA setting in this analysis, we keep to the strategy of a wide range of possible settings to ensure robustness of our findings. Within this work, we present results for two specific window lengths, namely 61 and 121 days. For the robustness check, we also tested other variants ranging from one month to two years, namely 21, 241, and 481 days, leading to comparable results.<sup>24</sup> Instead of usual ‘historical’ MA taking into account only the past information for given time, we use the ‘centred’ MA taking into account the same number of observation back as ahead. Both MA versions were analysed and found to produce to a large extent comparable results. The centred MA is therefore suggested to reduce the delay of the information flow. Moreover, the centred MA incorporates a convenient property that the price converges to it by definition that is exactly a feature one would expect from the fundamental value. Although undoubtedly our fundamental price approximation differ from the true fundamental value, the MA filter produces a series of an anticipated structure as depicted in Figures 12 and 13 in Appendix 2 for S&P500 and in Figures 14 and 15, 16 and 17, 18 and 19, 20 and 21, 22 and 23 for other respective indices.

### 5.3. Data description

We use daily closing prices of six world stock market indices as the base of our empirical dataset. For S&P500, we retrieve the closing prices of the index using the *Wolfram Mathematica* `FinancialData` function covering the period from 1994/02/23 to 2013/12/31, i.e. 5000 observations in total. For other indices, only the starting dates of the dataset vary by reason of different public holidays a calendar configurations around the world, i.e. 1994/04/22 for DAX,<sup>25</sup> 1994/11/02 for FTSE, 1994/02/23 for NASDAQ, 1993/09/03 for NIKKEI 225, and 1994/06/13 HSI. For each index we nonetheless also obtain comparable amount of 5000 observations with the same end date 2013/12/31. The fundamental price is simultaneously calculated as the centred MA as described in Subsection 5.2 and subtracted from the actual price following Equation 9.<sup>26</sup> Thus we obtain deviations  $x_t$  from the fundamental price that are the subject of further estimation. The span of the data is represented in Figure 12 and 13 where the original time series of prices  $p_t$  is depicted in the (a) part of the Figure together with the fundamental price  $p_t^*$  approximation and the (b) and (c) parts depict the implied series of deviations from the fundamental price  $x_t = p_t - p_t^*$ . Descriptive

<sup>24</sup>Results of this robustness testing are available upon request from authors.

<sup>25</sup>For DAX the data are available from 1990/11/26, i.e. 5850 observations till 2013/12/31.

<sup>26</sup>For the calculation of the fundamental price we need some extra data points preceding and succeeding the defined period, the complete dataset retrieved and used therefore consists of ‘4999 + MA window length’ observations.



Table 6: Descriptive statistics of empirical  $x_t$  time series

Data, MA period	Mean.	Med.	Min.	Max.	SD	Skew.	Kurt.	LQ	HQ	AC	AC $x_t^2$
SP500, 61	-.03	2.0	-145.2	113.0	26.5	-.65	5.5	-61.4	50.7	.87	.73
NASDAQ, 61	-.10	2.4	-753.5	639.0	81.8	.03	13.6	-170.7	147.9	.89	.81
DAX, 61	-.28	10.1	-939.6	716.1	163.2	-.50	5.4	-361.8	311.6	.89	.80
FTSE, 61	-.12	7.5	-702.5	404.6	122.7	-.58	5.1	-275.4	235.7	.88	.73
HSI, 61	.35	24.3	-4253.2	3463.5	562.2	-.29	6.7	-1186.7	1130.0	.90	.74
NIKKEI 225, 61	.16	6.8	-2249.8	1900.2	434.7	-.30	4.2	-953.8	799.4	.89	.77
SP500, 241	-.60	2.8	-252.9	160.1	48.4	-.64	4.5	-112.7	85.7	.96	.90
NASDAQ, 241	-1.78	1.6	-756.4	1253.7	168.1	1.01	11.3	-350.9	322.0	.97	.96
DAX, 241	-3.35	1.4	-1531.3	1242.6	330.4	-.21	4.5	-728.0	669.1	.97	.94
FTSE, 241	-.73	11.9	-1072.4	721.0	210.6	-.56	4.4	-479.4	388.5	.96	.90
HSI, 241	2.26	-5.5	-6505.5	7099.5	1177.6	.18	5.7	-2424.1	2282.0	.98	.94
NIKKEI 225, 241	-6.38	-22.6	-3497.8	2872.4	860.0	-.20	3.3	-1821.4	1581.3	.97	.92

*Note:* Sample means, medians, minima, maxima, standard deviations (SD), skewnesses, kurtoses, 2.5% (LQ) and 97.5% (HQ) quantiles, and autocorrelations (AC) are reported. Figures are rounded to 1 or 2 decimal digits.

statistics of  $x_t$  series for all indices and two MA lengths for the fundamental value approximation are summarised in [Table 6](#).

#### 5.4. Static NPSMLE estimates

In the estimation, our main goal is to verify the HAM ability to describe world stock market data and whether we obtain estimates of a reasonable precision using the NPSMLE method. A special focus is also devoted to possible differences and similarities between particular indices. For reader's convenience we briefly repeat the structure of the estimated model model:

$$Rx_t = \sum_{h=1}^H n_{h,t}(g_h x_{t-1} + b_h) + \epsilon_t, \quad (31)$$

$$n_{h,t} = \frac{\exp(\beta U_{h,t-1})}{\sum_{h=1}^H \exp(\beta U_{h,t-1})}, \quad (32)$$

$$U_{h,t-1} = (x_{t-1} - Rx_{t-2}) \frac{g_h x_{t-3} + b_h - Rx_{t-2}}{a\sigma^2}, \quad (33)$$

where  $\epsilon_t$  is an i.i.d. noise term sequence with normal distribution  $\mathcal{N}(0, sd^2)$ . Please, consult details of the model setting in [Subsection 4.1](#). To recap, in the 2-type model ( $H = 2$ ) we simultaneously estimate parameters  $\beta$ ,  $g_2$ ,  $b_2$ , and the *noise intensity*. In the 3-type model ( $H = 3$ ) we simultaneously estimate parameters  $\beta$ ,  $g_2$ ,  $g_3$ , and the *noise intensity*.

##### 5.4.1. Full sample estimates of the 2-type model

We start with the full sample static estimation summarised in [Table 7](#). Generally, we can observe broad similarities across all indices and markedly statistically significant estimates of a positive belief parameter  $g_2$  revealing superiority of trend following over contrarian strategies on markets. In contrast, the estimates of the intensity of choice—the switching parameter  $\beta$ —and the bias parameter  $b_2$  are largely statistically insignificant. While for the bias this is an expected result as there is no obvious reason why the trend following strategies should be somehow biased in the long-term, the insignificance of  $\hat{\beta}$  is an important and interesting result. We thus contrast a

Table 7: Empirical results of the 2-type  $\beta$  model estimation

Data, MA period	(a) $\hat{\beta}$			(b) $\hat{g}_2$			(c) $\hat{b}_2$			(d) $\widehat{noise\ intensity}$			(e) $LL$		
	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD
SP500, 61	.015	.040	.122	1.567	1.587	.233	.009	.003	.121	.653	.656	.108	-1.486	-1.491	.074
NASDAQ, 61	.002	.006	.146	1.717	1.715	.166	-.005	-.003	.092	.609	.609	.079	.117	.115	.038
DAX, 61	.001	.018	.112	1.640	1.646	.215	.008	.001	.117	.590	.601	.099	-1.259	-1.264	.081
FTSE, 61	-.000	.008	.113	1.671	1.668	.201	.001	.004	.117	.597	.602	.092	-.988	-.995	.059
HSI, 61	-.004	-.002	.150	1.724	1.727	.164	-.003	-.001	.098	.566	.570	.069	-.145	-.147	.036
N225, 61	.008	.021	.100	1.601	1.619	.210	.011	.007	.121	.593	.601	.103	-1.544	-1.546	.085
SP500, 241	-.021	-.009	.195	1.878	1.875	.127	.007	.001	.111	.388	.397	.064	.029	.019	.074
NASDAQ, 241	.038	.047	.189	1.882	1.869	.152	-.009	-.003	.105	.423	.419	.078	.422	.424	.111
DAX, 241	.014	.012	.247	1.932	1.928	.109	.007	.003	.103	.292	.315	.070	.328	.309	.096
FTSE, 241	.010	.012	.162	1.871	1.858	.143	.008	.004	.114	.408	.409	.071	-.153	-.157	.085
HSI, 241	.008	.006	.204	1.914	1.908	.130	.004	.002	.103	.351	.366	.080	.369	.363	.118
N225, 241	.003	.012	.147	1.865	1.850	.165	-.000	-.000	.125	.394	.392	.101	-.793	-.793	.160
Robustness check															
SP500 monthly, 13	-.016	-.031	.272	.886	.891	.238	.006	.002	.121	.863	.869	.104	-3.660	-3.666	.054
SP500 weekly, 13	-.085	-.103	.228	.953	.971	.224	-.003	-.001	.121	.865	.869	.098	-2.336	-2.349	.057
SP500 weekly, 49	.044	.089	.146	1.119	1.168	.241	.005	.002	.118	.724	.720	.103	-2.626	-2.615	.073
SP500 R=1.001, 61	.016	.035	.112	1.585	1.607	.241	.004	.004	.121	.656	.654	.110	-1.480	-1.485	.073
SP500 R=1.001, 241	.001	-.001	.196	1.889	1.880	.134	.005	.002	.111	.393	.397	.062	.029	.019	.073
SP500 $m_h=40$ , 61	-.032	-.041	.149	1.659	1.673	.173	-.002	-.002	.114	.584	.587	.075	-1.393	-1.402	.038
SP500 $m_h=40$ , 241	-.028	-.020	.298	1.908	1.905	.107	-.004	-.003	.099	.338	.355	.056	.100	.086	.054
SP500 $m_h=80$ , 61	-.047	-.054	.177	1.644	1.664	.182	.003	.003	.115	.573	.576	.065	-1.385	-1.392	.032
SP500 $m_h=80$ , 241	-.036	-.014	.309	1.908	1.906	.105	-.002	-.004	.097	.335	.352	.054	.106	.091	.056

*Note:* Results are based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians, means, and standard deviations (SD) are reported. ‘LL’ denotes log-likelihoods of estimated models representing statistical fits. Figures are rounded to 3 decimal digits.



large subpart of the HAM estimation literature (see [Section 2](#)) but confirm the main results of e.g. [Westerhoff and Reitz \(2005\)](#); [Boswijk et al. \(2007\)](#); [de Jong et al. \(2009b\)](#); [ter Ellen et al. \(2013\)](#); [Bolt et al. \(2014\)](#). Since the heterogeneity in trading regimes is confirmed by the significance of  $g_2$ , this might not be worrying as discussed in [Boswijk et al. \(2007, pg. 1995\)](#) or [Hommes \(2013, pg. 203\)](#) who emphasise that “this is a common result in non-linear switching regression models, where the parameter in the transition function is difficult to estimate and has a large standard deviation, because relatively large changes in  $\beta^*$  cause only small variation of the fraction  $n_t$ ”. [Teräsvirta \(1994\)](#) argues that this should not be worrying as long as there is significant heterogeneity in the estimated regimes.” Furthermore, as [Huisman et al. \(2010, pg. 17, 20\)](#) point out, “the significance of the intensity of choice is not a necessary condition for the switching to have added value to the fit of the model” and “the non-significant intensity of choice, as this indicates that the switching does not occur systematically”. The magnitudes of trend parameter estimates  $\hat{g}_2$ , that keep roughly between 1.6 and 1.9, might seem large, but it is important to note that they influence the price change only from circa 50% implied by the insignificance of the intensity of choice  $\beta$  keeping the population ratio of the two strategies stable around 0.5/0.5.

It is now important to contrast empirical findings with simulation results of [Section 4](#). Based on the analysis of the confidence bands in [Figure 7](#), our computational setting based on 1000 random runs, number of observations  $t = 5000$ , and the kernel approximation precision  $N = 500$  provides us with a reasonably high precision of the intensity of choice  $\beta$  estimates. [Figure 7](#) (bottom part) and further quantitative analysis clearly show that even for a very small  $\beta = 0.5$ , when estimating a time series of 5000 observations and considering 5% significance level,  $\beta$  is markedly statistically significant. Thus, if there is some behavioural switching present in our empirical data, it should have been detectable under similarly robust setting.

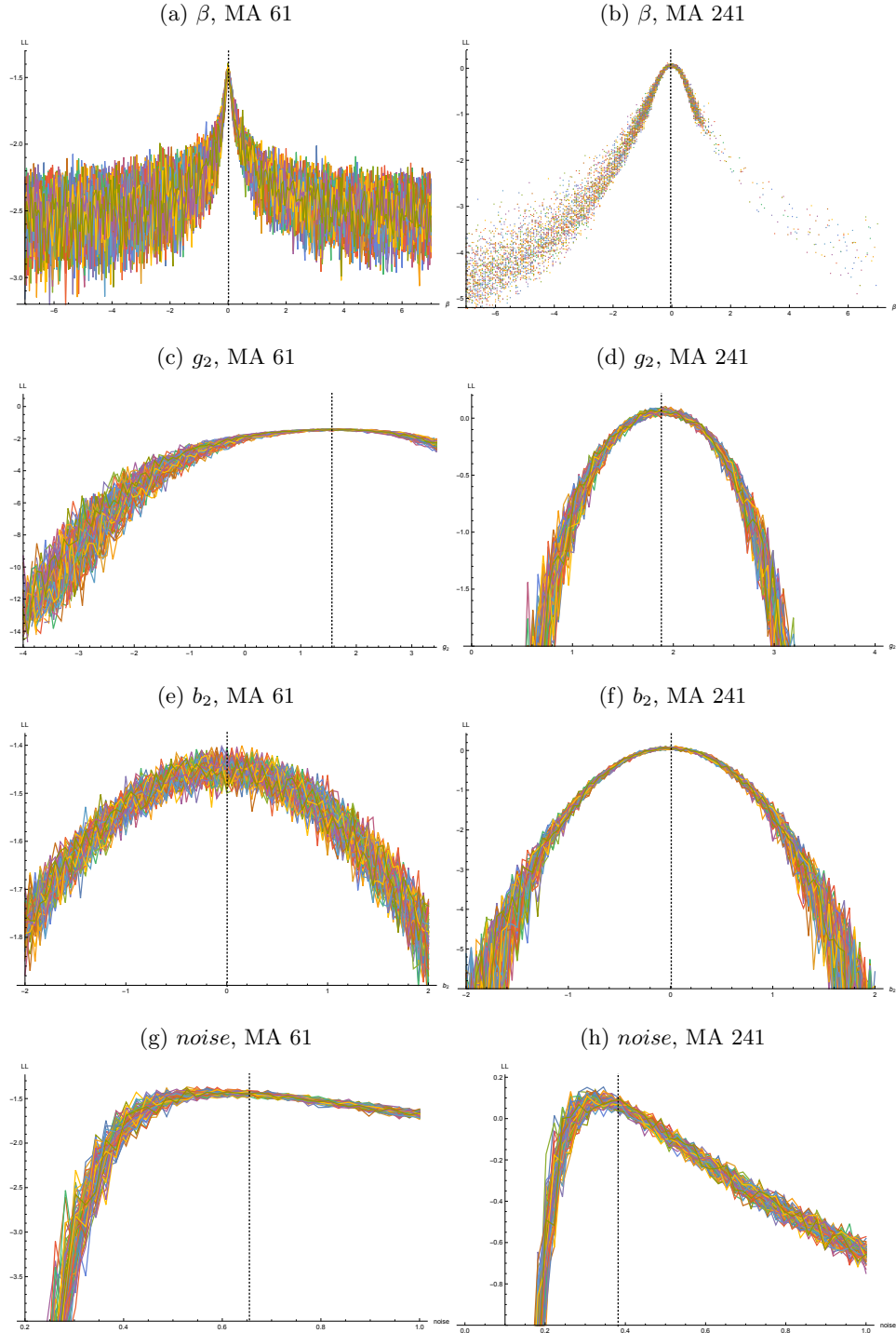
Differences across markets can be partly seen in the (d) column of [Table 7](#) between well efficiently estimated values of the *noise intensity*. Although the difference are often on the border of statistical significance, it might be worth mentioning that the highest stock market noise intensity has been estimated for the U.S. indices, specifically S&P500 in case of MA61 based fundamental value and NASDAQ in case of MA241 based fundamental value. Conversely, the lowest values has been estimated for DAX and the difference is circa 30% in case of the MA241 based fundamental value.

The level differences in values between the upper part of [Table 7](#) depicting results for the MA61 fundamental price approximation and the middle part with results for the MA241 is perhaps mainly the technical feature of different MA windows. It is therefore important to consider absolute values of estimated coefficient with this respect and compare both versions. Nevertheless, the main results concerning the positive sign and statistical significance of  $\hat{g}_2$  and insignificance of  $\hat{\beta}$  and  $\hat{b}_2$  keep similar as well as the main detected relative relationships between values of the *noise intensity*. Our most important results thus demonstrate robustness w.r.t. the choice of the fundamental value specification. The lower values of the *noise intensity* might be explained by reason of a better fundamental value approximation using bigger MA window.

#### 5.4.2. Behaviour of the simulated log-likelihood function

We verify the smoothness conditions and unique maxima presence of the simulated log-likelihood functions that are crucial for the parameter detection and identification also within the empirical application. However, the 5D surface resulting from simultaneous estimation of four parameters makes the graphical demonstration even more complicated than in [Subsection 5.4.2](#).

Figure 5: Simulated sub-log-likelihood fcn. for single parameters



*Note:* Results based on 100 random runs, S&P500 data, given MA fundamental price approximation, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ .  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distribution. Black dotted vertical lines depict estimated parameters (see [Table 7](#)).

Therefore, we again depict sub-log-likelihood functions in 2D and 3D, assuming other parameters fixed at estimated values from [Table 7](#).

[Figure 5](#) demonstrates partial 2D shapes of the simulated sub-log-likelihood function in direction of individual parameters. Generally we observe a bit rough shape in detail, but very consistent performance of the estimation method over all 100 random runs leading to unique maxima consistent with the full sample estimates in all cases. The shape is affected by the structure of the real world data which is far away from the regularity of the simulated dataset. For this reason, we also adapted the computational algorithm by increasing the number of initial points (see details in [Subsection 5.1](#)) so that it is able to deal well with the not-completely-smooth surface of the log-likelihood function. [Figure 24](#) in [Appendix 2](#) then visualises 3D simulated sub-functions for all combinations of four estimated parameters, keeping the other two fixed. The relative smoothness of the surface and well detectable regions of possible maxima keep generally retained corresponding to simulation results from [Subsection 4.3.8](#). For more ‘extreme’ combinations of parameters the model is again numerically unstable and we do not depict the surface for these regions. Equivalently to full sample estimation results, parameters  $\beta$  and  $g_2$  seem to be well detectable, while in the  $b_2$  direction [subfigures (b), (d), and (f)] the surface is very flat. Interestingly, based on visual inspection of subfigures (g) and (h) we suspect a small potential upward bias for the *noise intensity* estimates.

#### 5.4.3. Robustness check of the 2-type model

For the robustness check of the validity of estimated values (results are reported in [Table 7](#), bottom part), we not only use more than single MA specification of the fundamental value, but also consider several modification of the setup and even different frequency of the data. Equipped with the knowledge from previous analysis, we again only compute results for S&P500. Aside utilisation of weekly and monthly data, we also follow the robustness testing from [Section 4](#) and estimate the model using

1. 10 times higher assumed market risk free rate  $R = 1.001$ ;
2. nontrivial memory of agents defined via parameters  $m_h = \{40, 80\} \forall h$ .<sup>27</sup>

Three new dataset cover S&P500 weekly data from 1994/02/28 to 2013/12/30 (i.e. 1035 observations) and monthly data from 1994/03/01 to 2013/12/02 (i.e. 238 observations). Selected periods cover the same span as the original daily dataset. The MA lengths have been selected so that they resemble most closely the 61 and 241 days for the fundamental value specification, i.e. 13 and 49 weeks in case of weekly data and 13 months for monthly data. The assumed market risk free rate has been adjusted to reflect the modified data periodicity, namely to  $R = 1.0005$  for weekly data and  $R = 1.002$  for monthly data.

---

<sup>27</sup>Memory process is a substantial modification of the model structure. We employ the similar approach as [Barunik et al. \(2009\)](#), [Vacha et al. \(2012\)](#), and [Kukacka and Barunik \(2013\)](#), i.e. [Equation 29](#) is extended via memory parameters  $m_h$ :

$$U_{h,t-1} = \frac{1}{m_h} \sum_{l=0}^{m_h-1} \left[ (x_{t-1-l} - Rx_{t-2-l}) \frac{g_h x_{t-3-l} + b_h - Rx_{t-2-l}}{a\sigma^2} \right]. \quad (34)$$

The memory for each individual strategy is then randomly generated from the uniform distribution  $U(0, m_h)$ , therefore the average memory length resembles circa the one- or two-month period, i.e. 20 or 40 days.

The most important findings of the original empirical analysis of six world stock market indices remain unaffected under the robustness burden. The  $\hat{\beta}$  parameter still reveals evident statistical insignificance, the same does the bias parameter  $\hat{b}_2$ . Differences are basically observable at the level of trend parameters  $\hat{g}_2$  and *noise intensity*, but the behaviour keeps patterns uncovered within the original analysis— $\hat{g}_2$  slightly increases and *noise intensity* decreases moving from MA61 to MA241 fundamental value approximation. Results based on monthly and weekly data show considerably lower  $\hat{g}_2$ , the values fall even under 1 for the MA13 fundamental value specification—instead of strong trend chasing strategy only the weak trend chasing strategy is detected—but this seems to be again an implied technical side-effect of a small MA window that produces more average dynamics of the price deviations series. Monthly data with insufficient length of the estimated series (238 observations) expectedly, based on findings of the NPSMLE method performance analysis in [Section 4](#), perform the worst statistical fit compared to weekly and daily dataset. *Memory* only slightly increases the model fit and as the interconnected effect decreases the *noise intensity*. Nothing surprising is therefore found within the validity check of original results which prove robust to various data frequency specifications as well as modifications of interesting parameters in the model.

#### 5.4.4. Full sample estimates of the 3-type model

Estimation of a more-flexible 3-type model reveals markedly similar big picture as the estimation of the 2-type model. For the matter of computational time, based on our knowledge from the 2-type model estimation revealing large similarities across all estimated stock market indices we again only compute results for S&P500. Estimated parameters are reported in [Table 8](#). The only new conclusion is a statistical insignificance of the contrarian strategy represented by coefficient  $\hat{g}_3$  (exactly specified in the model and defined via the constraint  $g_3 < 0$ ). Although point estimates reported via median and mean values are negative, this is only an effect of the enforced  $g_3 < 0$  constraint. The distribution mass of estimates from all 500 runs concentrates close to 0 as depicted in [Figure 25](#) in [Appendix 2](#). The optimised function is likely to be very flat in the dimension of  $g_3$  parameter because the effect of a very weak contrarian strategy is overshadowed if combined with a very strong trend following strategy. The estimate of the intensity of choice  $\beta$  keeps its statistical insignificance and the trend following strategy coefficient  $\hat{g}_2$  retains its positive sign as well as high statistical significance. The absolute value of  $\hat{g}_2$  is naturally higher because the trend following strategy impacts the price via only the 1/3 weight in the 3-type model compared to 1/2 weight in the 2-type model (in both cases conditional on insignificant  $\hat{\beta}$ ). Taking those weights into account, we obtain very similar impact of the trend following strategy in both models. Comparing results for the MA61 and MA241 fundamental price approximation shows the very same effects as within the 2-type estimation, under MA241 we reveal somewhat stronger trend following strategy and lower intensity of stochastic noise.

#### 5.5. Rolling NPSMLE estimates

To confirm the robustness of the full sample static estimates in [Subsection 5.4](#) over time, we further investigate how the HAM estimation results might possibly change between 1994 and 2014. Utilising almost 20 years of data (5000 observations) used for the static estimates, we now estimate the HAM on one year (240 days) rolling samples with steps of two months (40 days). Based on results of the simulation Monte Carlo study in [Section 4](#), the one year period still represents relatively sufficient length of the estimated time series for a reasonable statistical inference based on the NPSMLE method. At the same time it is a relatively short period to detect

Table 8: Empirical results of the 3-type  $\beta$  model estimation

Data, MA p.	(a) $\hat{\beta}$		(b) $\hat{g}_2$		(c) $\hat{g}_3$		(d) $\widehat{noise\ intensity}$		(e) $LL$	
	Med.	SD	Med.	SD	Med.	SD	Med.	SD	Med.	SD
SP500, 61	-.003	.082	2.502	.175	-.123	.111	.550	.047	-.127	.022
SP500, 241	.007	.050	2.674	.217	-.032	.142	.403	.045	-.289	.094

*Note:* Results are based on 500 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians and standard deviations (SD) are reported. ‘LL’ denotes log-likelihoods of estimated models representing statistical fits. Figures are rounded to 3 decimal digits.

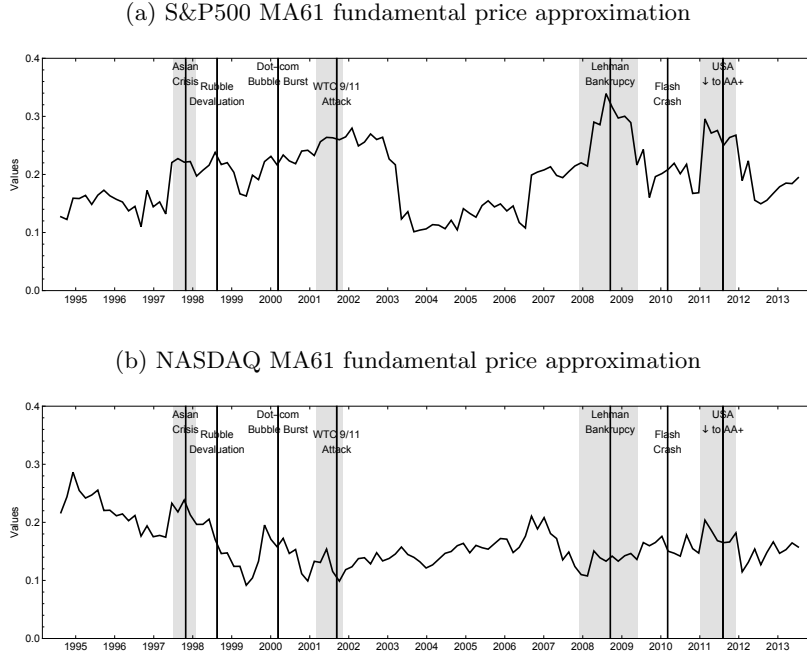
possible structural breaks in the data.<sup>28</sup> Due to high computational burden we have significantly decreased the robustness of the algorithm setting: we display results based on 200 runs and the number of initial points for the numerical optimisation has been decreased to 4. We also always discard only 10 observations as the initial stabilising periods. The ‘cost’ of such relaxation of the computational setting is reflected in the lower efficiency of estimates and the standard deviations of rolling estimates are expected higher compared to full sample static estimates. On the other hand, the rolling analysis still provides clear insight into the model dynamics and credible conclusion.

#### 5.5.1. Rolling estimates of the 2-type model

We depict rolling estimation results of the S&P500 in Figure 27 in Appendix 3. Rolling estimate results for other indices are reported in Figures 28, 29, 30, 31, and 32. Interpreting primarily results in Figure 27, we can observe relatively stable behaviour of the model throughout the entire investigated period. The  $\hat{g}_2$  and  $\widehat{noise\ intensity}$  estimates keep steadily around their long-term static estimates and exhibit high statistical significance as traceable in the (c) and (d) subparts of the Figure depicting rolling standard deviation for all coefficients. The  $\hat{\beta}$  and  $\hat{b}_2$  keep to zero and are statistically insignificant all the time. This is something we generally expect to observe as the full sample  $\hat{\beta}$  estimates (see Table 7) are close to zero, hence the dynamics is restrained and the model in fact boils down to a simple model which we further analyse in Subsection 5.6. Nonetheless, we might detect some signs of dynamics of  $\hat{g}_2$  e.g. around the Lehman Brothers bankruptcy and related U.S. recession between December 2007 and June 2009 in the MA61 case, but these shifts are strongly below the level of statistical significance and also largely dependent on the window length of the MA fundamental price approximation as one can see when comparing the (a) and (b) part of Figure 27. Some slight dynamics is also detectable for  $\widehat{noise\ intensity}$  which slightly increases in turbulent periods. The only clear dynamics seems to be observable at the level of the log-likelihood  $LL$ . With this respect we highlight the fact that a direct comparison of rolling log-likelihoods is methodologically disputable because it is based on different rolling sub-samples. However, we argue that since rolling datasets keep the same length and overlap by circa 83% between adjacent steps, the overall evolution of the  $LL$  pattern provides us with a valuable information. The fit is relatively higher during rather tranquil periods (e.g. at the beginning and the end of the sample period or between 2003 and 2007) and on the contrary it generally decreases during volatile periods (well detectable in Figures 12 and 13), especially during highlighted recessions. This might seem puzzling as during the high volatile periods the trend

<sup>28</sup>Various combinations of rolling sample windows and steps had been used in the preliminary analysis without impacting the overall results, e.g. comparing one month and one-half year steps. The outcomes of the preliminary analysis are available from authors upon request.

Figure 6: Rolling behaviour of the SD of the  $\hat{g}_2$  estimate I.



*Note:* Bold black full line depicts standard deviation of the  $\hat{g}_2$  estimate. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

following strategy is likely to attract attention of market participants, however such behaviour is in accord with the overall stability of the rolling estimates close to long-term values.

A potential dynamics is, however, much more observable at the level of standard deviations [see detached [Figure 6](#) and subparts (c) and (d) of [Figure 27](#)]. The standard deviation of the trend following coefficient  $g_2$  clearly jumps up in volatile periods such as recession around the WTC 9/11 attack, Lehman Bankruptcy, or downgrade of USA ranking to AA+. This can be interpreted as a sign of increased presence of contrarians (nonetheless, still being a large minority) during such turbulent stock market periods. Although the method faces difficulties to adjust the average absolute value of the trend-following coefficient  $\hat{g}_2$  over 500 repeated runs, it detects increasing population of contrarians via less efficient estimates of the effect of trend followers. A similar pattern is observable for the *noise intensity* and *LL*. This is again nothing surprising as econometrics models generally perform better in periods of market stability. Using larger MA241 fundamental value approximation naturally decreases the flexibility of the estimation to detect effects of single events and so the captured dynamics is considerably more stable.

We also observe some interesting signs of a specific and economically well interpretable dynamic behaviour for other indices illustrated in [Figure 33](#) in [Appendix 2](#). For NASDAQ behaviour, although otherwise considerably more stable compared to S&P500, the worth mentioning is especially the drop of the standard deviation of  $\hat{\beta}$  and increase of the standard deviation of the trend parameter  $\hat{g}_2$  around the Asian Crisis in 1997 and the Dot-com Bubble Burst in 2000. Since



NASDAQ is especially used for trading technological and IT companies, this makes somewhat sense as technological companies are often based or produce in Asia and IT companies were hit by the Dot-com Bubble much more than other sectors. Similar behaviour is not observed for the more general S&P500. For both European indices DAX and FTSE the model on the level of standard deviations does reflect the crisis around the Rubble devaluation in 1998, again not captured by S&P500. The effect of both worldwide crises in 2001 and 2008-2009 is also captured, but not the Asian Crisis of 1997. Interestingly, the HSI data representing Asia captures the effect of 1997 Asian Crisis as well as the worldwide crises of 2008-2009 and 2011, but does not reflect neither the 1998 Rubble devaluation, nor WTC 2001/9/11 attack. Behaviour of model under NIKKEI 225 data resembles much more the patterns observed for S&P500 than for HSI. All this can be attributed to increased presence of contrarians at specific periods on specific markets detected via less efficient estimates of the effect of trend followers. Finally for some indices, namely NASDAQ, FTSE and HSI (but not for the other three) we can observe some intriguing negative correlation between standard deviations of  $\hat{\beta}$  and  $\hat{g}_2$  in some turbulent periods.

In general, although we can reveal some patterns of interpretable dynamics for individual indices, the rolling approach in particular strongly supports the stability of model behaviour over time and thus confirms the validity of full sample estimation results from [Subsection 5.4.1](#).

#### 5.6. Estimation of market fractions

Our findings from the performed estimations in [Subsection 5.4](#) and [Subsection 5.5](#), mainly the overall statistical insignificance of the intensity of choice  $\hat{\beta}$ , statistically insignificant sizes of  $\hat{g}_3$  coefficients of contrarian strategies, and stability of rolling coefficients, lead us to another, this once truly significant modification of the model. Interpreting these results, hitherto model specifications do not seem to correspond to the data fully. In the 2-type model, the insignificant  $\hat{\beta}$  coefficient implicates stable population ratio of trading strategies  $n_{1,t}/n_{2,t} \doteq 0.5/0.5$ , which means that the population of fundamentalists is forced to be of almost the same magnitude as the population of chartists throughout the entire span of the dataset. Thus the model in fact boils down to a simple weighted AR(1) process and different types of traders cannot be identified because they do not switch over time. In such a case the trend and bias parameters  $\hat{g}_2$  and  $\hat{b}_2$  (or  $\hat{g}_3$  in the 3-type model) can be viewed as nuisance parameters—they to a large extent lose the original model interpretation and we cannot fully trust the estimated magnitudes of these parameters. Although we understand that it is generally very complicated for the estimation method to detect some systematic evolutionary switching between trading strategies when it is exposed to the full dataset (and therefore the average zero  $\hat{\beta}$  coefficient seems reasonable), we cannot agree with such a strong assumption of similar population magnitudes for both strategies. Moreover, contrarians in the 3-type model, who technically account for 1/3 of the population size as the  $\hat{\beta}$  coefficient is steadily insignificant, in fact behave as fundamentalists in terms of their price impact because sizes of  $\hat{g}_3$  coefficients are small and statistically insignificant. These findings from the analysis of the 2-type and 3-type model estimation imply two important conclusions. First, the 3-type model does not really help us to capture additional features of the data-generating process and rather deviates the implied market fraction. Second, it suggests there might be more fundamentalists than chartists on real markets and therefore the almost fixed population ratio of trading strategies  $n_{1,t}/n_{2,t} \doteq 0.5/0.5$  as the result of the 2-type model estimation is much likely not capturing the real market population proportions.

Therefore, as a consequence of previous findings, we trivialise the simulated model ([Equation 27](#), [Equation 28](#), and [Equation 29](#)) via disabling the evolutionary switching behaviour and



fixing the population ratio of trading strategies to  $n_{1,t}/n_{2,t} = \text{const.}$  A rationale of this step is further supported by overall stability of rolling estimates. Equation 28 and Equation 29 are now replaced by Equation 36 and the coefficient  $n_1$ , which we further call percentage *fraction* of fundamentalists, is to be estimated instead of the switching coefficient  $\beta$ :

$$Rx_t = \sum_{h=1}^H n_h f_{h,t} + \epsilon_t \equiv \sum_{h=1}^H n_h (g_h x_{t-1} + b_h) + \epsilon_t, \quad (35)$$

$$n_1 = 1 - n_2, \quad (36)$$

where  $H = 2$  in the 2-type model. Interval for the starting points random generation is constrained to  $\langle 0.3, 0.9 \rangle$  for *fraction* and to  $\langle 1.5, 2.5 \rangle$  for  $g_2$ , the other setting remains the same as in Subsection 5.4. The modified setup keeps the logic of aforementioned findings and does not distract the structure of the original model. On the other hand, the population ratio of trading strategies  $n_1/n_2$  and implied percentage *fraction* of fundamentalists on the market is now a direct subject of the interest.

#### 5.6.1. Full sample estimates of the 2-type fraction model

Outcomes of the full sample static estimation of all six stock market indices are reported in Table 24 in Appendix 1. The main interest lies in the behaviour of the new variable *fraction* representing the percentage market fraction of fundamentalists ( $g_1 = b_1 = 0$ ). All other variables behave at average very similar as in the 2-type  $\beta$  model estimation, moreover we do not longer observe considerable distinctions caused by the MA window length for the fundamental value approximation.

The *fraction* coefficient is strongly statistically significant with its value closely around 0.56, leaving only 44% of the market population to chartistic strategies. The model therefore suggests overall proportional dominance of the fundamental strategy on all investigated world stock markets. Estimates of the trend following coefficient  $g_2$  are generally higher compared to values for the 2-type  $\beta$  model estimation reported in Table 24 but one must realise that within the 2-type *fraction* model the trend following strategy is relatively weaker in terms of impact to the market price (see Equation 13) because the proportion of these strong trend chasers is lower than 0.5. If we consider market proportions incorrectly implied by the 2-type  $\hat{\beta}$  coefficient and related  $\hat{g}_2$  and compare it to *fraction* and related  $\hat{g}_2$  estimated in this section according to Equation 13, we deduce almost similar impact. This confirms our suspicion about an improper specification of the model with insignificant  $\hat{\beta}$  and we corrected for this misspecification introducing *fraction* specification via Equation 36. Evolutionary switching between strategies can be now captured via changes in the *fraction* coefficient in its smooth form using the rolling approach as asserted by Teräsvirta (1994, pg. 217): “if one assumes that the agents make only dichotomous decisions or change their behaviour discretely, it is unlikely that they do this simultaneously. Thus if only an aggregated process is observed, then the regime changes in that process may be more accurately described as being smooth rather than discrete.” Nonetheless, the rolling approach does not reveal any significant dynamics in the behaviour of *fraction* which again only confirms the validity of full sample estimation results from Subsection 5.6.

### 5.6.2. Behaviour of the simulated log-likelihood function

Conclusions for the smoothness conditions and unique maxima presence of the simulated log-likelihood functions for the  $\beta$  model (see Subsection 5.4.2) hold generally identically for the *fraction* model. Here we only depict sub-log-likelihood functions in 2D assuming other parameters fixed at estimated values from Table 24. In Figure 26 in Appendix 2 we demonstrate 2D shapes of the simulated sub-function in direction of individual parameters. Differences compared to the  $\beta$  model are threefold:

1. in the *fraction* direction [subfigures (a) and (b)] the function behaves more ‘nicely’;
2. in the  $b_2$  direction [subfigures (e) and (f)] we do not observe any optimum, thus the identification of the parameter seems problematic;
3. on the contrary, we do not longer suspect the potential upward bias for the *noise intensity* estimates [subfigures (g) and (h)] compared to Subsection 5.6.2.

### 5.6.3. Robustness check of the 2-type fraction model

We employ an identical (except for now irrelevant effect of memory) robustness check as in the previous case for the  $\beta$  model also for the *fraction* model. Results of the weekly and monthly data estimation and the model assuming higher market risk free rate are reported in Table 24 (bottom part). Basic conclusions for the robustness and validity check of the  $\beta$  model from Subsection 5.4.3 hold identically for the *fraction* model. The *fraction*,  $g_2$ , and *noise intensity* generally reveal strong statistical significance, the opposite does the bias parameter  $\hat{b}_2$ . Differences are again observable at the level of trend parameters  $\hat{g}_2$  and *noise intensity* based on monthly and weekly data—results show lower  $\hat{g}_2$  and higher *noise intensity* compared to daily data and monthly data perform the worst statistical fit compared to weekly and daily dataset. These findings are once again likely to be an implied technical side-effect of small MA window.

## 6. Conclusion

This paper proposes innovative computational framework for empirical estimation of FABMs. Motivated by the lack of general consensus on the estimation methodology, not many examples on structural estimation of FABMs, and inconclusive results in recent FABM literature, we introduce a general framework for estimation of full-fledged FABMs that significantly reduce the importance of restrictive theoretical assumptions. Because for many FABMs no closed-form representation of the likelihood function exists, we follow the Kristensen and Shin (2012) framework of a simulated MLE based on nonparametric kernel methods. In situations when we cannot derive the usual MLE, simulated MLE constitutes an opportune estimation method for general class of FABMs.

We customise the NPSMLE methodology of Kristensen and Shin (2012) and elaborate its capability for FABMs estimation purposes. To start with, we apply the methodology to the popular and widely analysed model of Brock and Hommes (1998). We extensively test small sample properties of the estimator via Monte Carlo simulations and confirm the ability of the NPSMLE method to reveal true parameters with high precision. We further show that theoretical properties of the estimator, the consistency and asymptotic efficiency, also hold in small samples for the model. Next, we assess the impact of the stochastic noise intensity in the system and investigate the robustness of the estimation method w.r.t. various modifications of the model as well as of the estimation algorithm. Finally, using graphical computational tools we analyse behaviour of simulated log-likelihood functions. Based on generally very smooth shape with a unique maxima

we assume that the regularity conditions are met for the HAM and the identification of parameters is assured.

We present the full sample static estimates to reveal robust average relationships in the 2-type and the 3-type [Brock and Hommes \(1998\)](#) model using cross section of world stock markets as well as rolling window approach to detect possible dynamics in the behaviour of market coefficients over time and eventual structural breaks. The crucial result of our analysis is the statistical insignificance of the switching coefficient  $\hat{\beta}$ . This is a common result in the existing literature, but on the other hand we contrast another part of the HAM estimation literature reporting significant  $\hat{\beta}s$  for various specific markets. In contrary, our estimation results of the 2-type model reveal markedly statistically significant belief parameters defining heterogeneous trading regimes with an absolute superiority of trend-following over contrarian strategies. Our findings further indicate robustness w.r.t. the fundamental value specification and remain largely unaffected under the robustness burden of different than daily data frequency, jumps in market risk free rate, or introducing of agents' memory. Graphical inspection of simulated log-likelihood functions reveals a bit rough surface, but very consistent performance of the estimation method over all random runs leading to unique maxima. The adapted computational algorithm is, however, able to deal well with the not-completely-smooth surface of the simulated log-likelihood function and the important identification feature is thus verified also for the empirical application.

Both main results are also stable over the entire period confirmed via rolling estimation approach which primarily supports the validity of the full sample static estimates. Interesting signs of a specific behaviour are detectable for rolling estimation of individual indices on the level of standard deviations. S&P500 data detects turbulent periods around the WTC 9/11 attack, Lehman Bankruptcy, or downgrade of the U.S. ranking to AA+ in 2011 but also other tracked world events. NASDAQ reflects mainly the Asian Crisis in 1997 and the Dot-com Bubble Burst in 2000 which is, however, not observed for a more general S&P500. European indices DAX and FTSE can detect the crisis around the Rubble devaluation in 1998 but not the Asian Crisis of 1997. HSI conversely captures the effect of 1997 Asian Crisis but not the 1998 Rubble devaluation. This can be interpreted as a sign of increased presence of contrarians during such turbulent stock market periods detected via less efficient estimates of trend following coefficients. Next, estimation of a more-flexible 3-type model with the mix of fundamental, trend following, and contrarian strategy further suggests redundancy of the contrarian strategy for the overall model fit.

Finally, we correct for possible model misspecification introducing fixed fraction of the fundamental strategy instead of switching coefficient  $\beta$  which is estimated statistically insignificant. The fraction is, however, the subject of empirical estimation and is found strongly statistically significant for all analysed indices. A strong trend chasing strategy is then expressed via trend coefficient around 2 for all indices in the *fraction* model. The magnitude of the fundamentalists' population closely around 56% represents overall proportional dominance of fundamentalists over trend following chartists on world stock markets and even stable in time.

## References

- Alfarano, S., T. Lux, and F. Wagner (2005). Estimation of agent-based models: The case of an asymmetric herding models. *Computational Economics* 26, 19–49.
- Alfarano, S., T. Lux, and F. Wagner (2006). Estimation of a simple agent-based model of financial markets: An application to australian stock and foreign exchange data. *Physica A: Statistical Mechanics and its Applications* 370(1), 38–42. Econophysics Colloquium Proceedings of the International Conference.
- Alfarano, S., T. Lux, and F. Wagner (2007). Empirical validation of stochastic models of interacting agents. *The European Physical Journal B* 55(2), 183–187.
- Allen, H. and M. P. Taylor (1990). Charts, noise and fundamentals in the london foreign exchange market. *The Economic Journal* 100(400), 49–59.
- Altissimo, F. and A. Mele (2009). Simulated non-parametric estimation of dynamic models. *The Review of Economic Studies* 76(2), 413–450.
- Amilon, H. (2008). Estimation of an adaptive stock market model with heterogeneous agents. *Journal of Empirical Finance* 15(2), 342–362.
- Andersen, T. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *The Journal of Finance* 57(3), 1239–1284.
- Barunik, J. and J. Kukacka (2015). Realizing stock market crashes: stochastic cusp catastrophe model of returns under time-varying volatility. *Quantitative Finance* 15(6), 959–973.
- Barunik, J., L. Vacha, and M. Vosvrda (2009). Smart predictors in the heterogeneous agent model. *Journal of Economic Interaction and Coordination* 4, 163–172.
- Barunik, J. and M. Vosvrda (2009). Can a stochastic cusp catastrophe model explain stock market crashes? *Journal of Economic Dynamics & Control* 33, 1824–1836.
- Bass, F. (1969). A new product growth for model consumer durables. *Management Science* 15, 215–227.
- Biondi, Y., P. Giannoccolo, and S. Galam (2012). Formation of share market prices under heterogeneous beliefs and common knowledge. *Physica A: Statistical Mechanics and its Applications* 391(22), 5532–5545.
- Bolt, W., M. Demertzis, C. Diks, and M. van der Leij (2011). Complex methods in economics: an example of behavioral heterogeneity in house prices. Technical report, De Nederlandsche Bank.
- Bolt, W., M. Demertzis, C. G. H. Diks, C. H. Hommes, and M. van der Leij (2014). Identifying booms and busts in house prices under heterogeneous expectations. Technical report, De Nederlandsche Bank Working Paper No. 450.
- Boswijk, H. P., C. H. Hommes, and S. Manzan (2007, February). Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics & Control* 31(2), 1938–1970.
- Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *The Economic Journal* 114(497), 592–621.
- Branch, W. A. and G. W. Evans (2006). Intrinsic heterogeneity in expectation formation. *Journal of Economic Theory* 127(1), 264–295.
- Brock, W., J. Lakonishok, and B. LeBaron (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance* 47(5), 1731–1764.
- Brock, W. A. and C. H. Hommes (1997). A rational route to randomness. *Econometrica* 65(5), 1059–1095.
- Brock, W. A. and C. H. Hommes (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics & Control* 22, 1235–1274.
- Browning, M., L. P. Hansen, and J. J. Heckman (1999). Chapter 8 micro data and general equilibrium models. In *Handbook of Macroeconomics*, Volume 1, Part A of *Handbook of Macroeconomics*, pp. 543–633. Elsevier.
- Canova, F., F. Schorfheide, and H. van Dijk (2014). Introduction to recent advances in methods and applications for DSGE models. *Journal of Applied Econometrics* 29(7), 1029–1030.
- Chen, S.-H., C.-L. Chang, and Y.-R. Du (2012, 6). Agent-based economic models and econometrics. *The Knowledge Engineering Review* 27, 187–219.
- Chen, Z. and T. Lux (2015). Estimation of sentiment effects in financial markets: A simulated method of moments approach. FinMaP-Working Paper 37, University of Kiel, Department of Economics, Kiel.
- Chiarella, C., R. Dieci, and X.-Z. He (2009). *Handbook of Financial Markets: Dynamics and Evolution*, Chapter 5: Heterogeneity, Market Mechanisms and Asset Price Dynamics, pp. 277–344. North-Holland, Elsevier, Inc., Amsterdam.
- Chiarella, C. and X.-Z. He (2002). An adaptive model on asset pricing and wealth dynamics with heterogeneous trading strategies. Technical report, University of Technology, Sydney, Australia.
- Chiarella, C., X.-Z. He, and C. Hommes (2006). A dynamic analysis of moving average rules. *Journal of Economic*

- Dynamics and Control* 30(9-10), 1729–1753. Computing in economics and finance 10th Annual Conference on Computing in Economics and Finance.
- Chiarella, C., X.-Z. He, and R. C. Zwickels (2014). Heterogeneous expectations in asset pricing: Empirical evidence from the SP500. *Journal of Economic Behavior & Organization* 105, 1–16.
- Chiarella, C., S. ter Ellen, X.-Z. He, and E. Wu (2015). Fear or fundamentals? heterogeneous beliefs in the european sovereign CDS markets. *Journal of Empirical Finance* 32, 19–34.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* 1(2), 223–236.
- Cont, R. (2007). Volatility clustering in financial markets: Empirical facts and agent-based models. In G. Teyssiere and A. Kirman (Eds.), *Long Memory in Economics*, pp. 289–309. Springer Berlin Heidelberg.
- Cornea, A., C. Hommes, and D. Massaro (2013). Behavioral heterogeneity in U.S. inflation dynamics. Tinbergen Institute Discussion Paper 13-015/II, Tinbergen Institute, Amsterdam and Rotterdam.
- Cox, J. C., J. Ingersoll, Jonathan E., and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica* 53(2), 385–407.
- De Grauwe, P. and M. Grimaldi (2005). Heterogeneity of agents, transactions costs and the exchange rate. *Journal of Economic Dynamics & Control* 29, 691–719.
- De Grauwe, P. and M. Grimaldi (2006). Exchange rate puzzles: A tale of switching attractors. *European Economic Review* 50, 1–33.
- de Jong, E., W. F. C. Verschoor, and R. C. J. Zwickels (2009a). Behavioural heterogeneity and shift-contagion: Evidence from the asian crisis. *Journal of Economic Dynamics & Control* 33, 1929–1944.
- de Jong, E., W. F. C. Verschoor, and R. C. J. Zwickels (2009b). A heterogeneous route to the european monetary system crisis. *Applied Economic Letters* 16, 929–932.
- de Jong, E., W. F. C. Verschoor, and R. C. J. Zwickels (2010). Heterogeneity of agents and exchange rate dynamics: Evidence from the EMS. *Journal of International Money and Finance* 29, 1652–1669.
- Diks, C. and R. Weide (2005). Herding, a-synchronous updating and heterogeneity in memory in a CBS. *Journal of Economic Dynamics & Control* 29, 741–763.
- Duan, J.-C. and J.-G. Simonato (1998). Empirical martingale simulation for asset prices. *Management Science* 44(9), 1218–1233.
- Ecemis, I., E. Bonabeau, and T. Ashburn (2005). Interactive estimation of agent-based financial markets models: Modularity and learning. In *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation, GECCO '05*, New York, NY, USA, pp. 1897–1904. ACM.
- Edgeworth, F. Y. (1881). *Mathematical psychics: An essay on the application of mathematics to the moral sciences*, Volume 10. Kegan Paul.
- Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review* 81(4), 922–930.
- Fagiolo, G., A. Moneta, and P. Windrum (2007). A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems. *Computational Economics* 30(3), 195–226.
- Fagiolo, G., M. Napoletano, and A. Roventini (2008). Are output growth-rate distributions fat-tailed? some evidence from oecd countries. *Journal of Applied Econometrics* 23(5), 639–669.
- Fagiolo, G. and A. Roventini (2012). Macroeconomic policy in dsge and agent-based models. *Revue de l'OFCE* 5(124), 67–116.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* 25(2), 383–417.
- Franke, R. (2009). Applying the method of simulated moments to estimate a small agent-based asset pricing model. *Journal of Empirical Finance* 16, 804–815.
- Franke, R. and F. Westerhoff (2011). Estimation of a structural stochastic volatility model of asset pricing. *Computational Economics* 38(1), 53–83.
- Franke, R. and F. Westerhoff (2012). Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. *Journal of Economic Dynamics and Control* 36(8), 1193–1211. Quantifying and Understanding Dysfunctions in Financial Markets.
- Frijns, B., T. Lehnert, and R. C. J. Zwickels (2010). Behavioral heterogeneity in the option market. *Journal of Economic Dynamics & Control* 34, 2273–2287.
- Gaunersdorfer, A. (2000). Adaptive beliefs and the volatility of asset prices. Technical report, Vienna University of Economics, Austria, Working Paper No. 74.
- Ghoshghadze, J. and T. Lux (2015). Bringing an elementary agent-based model to the data: Estimation via GMM and an application to forecasting of asset price volatility. FinMaP-Working Paper 38, University of Kiel, Department

- of Economics, Kiel.
- Gilli, M. and P. Winker (2003). A global optimization heuristic for estimating agent based models. *Computational Statistics & Data Analysis* 42, 299–312.
- Goldbaum, D. (1999). A nonparametric examination of market information: application to technical trading rules. *Journal of Empirical Finance* 6(1), 59–85.
- Goldbaum, D. and R. C. Zwinkels (2014). An empirical examination of heterogeneity and switching in foreign exchange markets. *Journal of Economic Behavior & Organization* 107, Part B, 667–684. Empirical Behavioral Finance.
- Gordon, M. J. (1962). The savings investment and valuation of a corporation. *The Review of Economics and Statistics* 44(1), 37–51.
- Grazzini, J. and M. Richiardi (2015). Estimation of ergodic agent-based models by simulated minimum distance. *Journal of Economic Dynamics and Control* 51, 148–165.
- Grazzini, J., M. Richiardi, and L. Sella (2013). Indirect estimation of agent-based models. an application to a simple diffusion model. *Complexity Economics* 1(2), 25–40.
- Hommes, C. (2013). *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press. Cambridge Books Online.
- Hommes, C. and D. Veld (2015). Booms, busts and behavioral heterogeneity in stock prices. Technical report, University of Amsterdam, CeNDEF Working Paper.
- Hommes, C. and F. O. O. Wagener (2009). *Handbook of Financial Markets: Dynamics and Evolution*, Chapter 4: Complex Evolutionary Systems in Behavioral Finance, pp. 217–276. North-Holland, Elsevier, Inc., Amsterdam.
- Hommes, C. H. (2006). Chapter 23: Heterogeneous agent models in economics and finance. In L. Tesfatsion and K. Judd (Eds.), *Handbook of Computational Economics*, Volume 2 of *Handbook of Computational Economics*, pp. 1109–1186. Elsevier.
- Howitt, P. (2012). What have central bankers learned from modern macroeconomic theory? *Journal of Macroeconomics* 34(1), 11–22. Has macro progressed?
- Huisman, R., R. A. F. Maliepaard, and R. C. J. Zwinkels (2010). Heterogeneous agents in electricity forward markets. Technical report, Erasmus University Rotterdam.
- Isakov, D. and M. Hollstein (1999). Application of simple technical trading rules to swiss stock prices: Is it profitable? Technical report, HEC-Universite de Geneve, Banque Cantonale de Geneve.
- Janssen, M. C. W. (1993). *Microfoundations: A Critical Inquiry*. Routledge.
- Jones, M. C., J. S. Marron, and S. J. Sheather (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association* 91(433), 401–407.
- Kahneman, D. and A. N. Tversky (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185, 1124–1131.
- Kahneman, D. and A. N. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47(2), 263–291.
- Kaldor, N. (1961). *Capital accumulation and economic growth*. Springer.
- Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money*. London: Macmillan.
- Kirman, A. (1991). *Money and Financial Markets*, Chapter Epidemics of Opinion and Speculative Bubbles in Financial Markets, pp. 354–368. Macmillan, New York, USA.
- Kouwenberg, R. and R. Zwinkels (2014). Forecasting the US housing market. *International Journal of Forecasting* 30(3), 415–425.
- Kouwenberg, R. and R. C. J. Zwinkels (2015, 06). Endogenous price bubbles in a multi-agent system of the housing market. *PLoS ONE* 10(6), e0129070.
- Kristensen, D. (2009, 10). Uniform convergence rates of kernel estimators with heterogeneous dependent data. *Econometric Theory* 25, 1433–1445.
- Kristensen, D. and Y. Shin (2012). Estimation of dynamic models with nonparametric simulated maximum likelihood. *Journal of Econometrics* 167(1), 76–94.
- Kukacka, J. and J. Barunik (2013). Behavioural breaks in the heterogeneous agent model: The impact of herding, overconfidence, and market sentiment. *Physica A: Statistical Mechanics and its Applications* 392(23), 5920–5938.
- LeBaron, B. and L. Tesfatsion (2008). Modeling macroeconomies as open-ended dynamic systems of interacting agents. *The American Economic Review* 98(2), 246–250.
- Lof, M. (0). Rational speculators, contrarians, and excess volatility. *Management Science* 0(0), null.
- Lof, M. (2012). Heterogeneity in stock prices: A STAR model with multivariate transition function. *Journal of Economic Dynamics and Control* 36(12), 1845–1854.
- Lucas, R. E. J. (1972). Expectations and the neutrality of money. *Journal of Economic Theory* 4(2), 103–124.

- Lucas, R. E. J. (1978). Asset prices in an exchange economy. *Econometrica* 46, 1429–1445.
- Lui, Y.-H. and D. Mole (1998). The use of fundamental and technical analyses by foreign exchange dealers: Hong Kong evidence. *Journal of International Money and Finance* 17(3), 535–545.
- Manzan, S. and F. H. Westerhoff (2007). Heterogeneous expectations, exchange rate dynamics and predictability. *Journal of Economic Behavior & Organization* 64, 111–128.
- Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica* 29(3), 315–335.
- Recchioni, M. C., G. Tedeschi, and M. Gallegati (2015). A calibration procedure for analyzing stock price dynamics in an agent-based framework. *Journal of Economic Dynamics & Control* 60, 1–25.
- Reitz, S. and U. Slopek (2009). Non-linear oil price dynamics: A tale of heterogeneous speculators? *German Economic Review* 10(3), 270–283.
- Reitz, S. and F. H. Westerhoff (2007). Commodity price cycles and heterogeneous speculators: A STAR-GARCH model. *Empirical Economics* 33, 231–244.
- Sargent, T. J. (1993). *Bounded Rationality in Macroeconomics*. Oxford: Clarendon Press.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19(3), 425–442.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics* 69(1), 99–118.
- Simon, H. A. (1957). *Models of Man*. New York: Wiley.
- Sullivan, R., A. Timmermann, and H. White (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance* 54(5), 1647–1691.
- Taylor, M. P. and H. Allen (1992). The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance* 11(3), 304–314.
- ter Ellen, S., W. F. Verschoor, and R. C. Zwinkels (2013). Dynamic expectation formation in the foreign exchange market. *Journal of International Money and Finance* 37, 75–97.
- ter Ellen, S. and R. C. Zwinkels (2010). Oil price dynamics: A behavioral finance approach with heterogeneous agents. *Energy Economics* 32(6), 1427–1434.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* 89(425), 208–218.
- Vacha, L., J. Barunik, and M. Vosvrda (2012). How do skilled traders change the structure of the market. *International Review of Financial Analysis* 23(0), 66–71. Complexity and Non-Linearities in Financial Markets: Perspectives from Econophysics.
- Verschoor, W. F. and R. C. Zwinkels (2013). Do foreign exchange fund managers behave like heterogeneous agents? *Quantitative Finance* 13(7), 1125–1134.
- Vigfusson, R. (1997). Switching between chartists and fundamentalists: A markov regime-switching approach. *International Journal of Finance and Economics* 2, 291–305.
- Wan, J.-Y. and C.-W. Kao (2009). Evidence on the contrarian trading in foreign exchange markets. *Economic Modelling* 26, 1420–1431.
- Westerhoff, F. H. and S. Reitz (2003). Nonlinearities and cyclical behavior: The role of chartists and fundamentalists. *Studies in Nonlinear Dynamics & Econometrics* 7(4), 1–13.
- Westerhoff, F. H. and S. Reitz (2005). Commodity price dynamics and the nonlinear market impact of technical traders: Empirical evidence for the us corn market. *Physica A* 349, 641–648.
- Winker, P. and M. Gilli (2001). Indirect estimation of the parameters of agent based models of financial markets. Technical report, International University in Germany and University of Geneva, Switzerland.
- Winker, P., M. Gilli, and V. Jeleskovic (2007). An objective function for simulation based inference on exchange rate data. *Journal of Economic Interaction and Coordination* 2(2), 125–145.



## **Appendix 1: Supplementary tables**

On the following pages, supplementary tables are provided.

Table 9: Estimation methods of FABMs II. b)

Models	#	Data	Type	Fit	IOC
Barunik and Vosvrda (2009)	8,17	d:1987–1988,2001–2002	s	$pseudo-R^2$ up to .8	-
Barunik and Kukacka (2015)	10	d:6739,409 o.	s	$pseudo-R^2=.8, .86$	-
Bolt et al. (2011)	5	q:164 o.	re	NA	2716(i),12420(i)
Bolt et al. (2014)	4	q:178 o.	re	NA	795(i)–26333(i)
Cornea et al. (2013)	2	q:204 o.	U.S. inflation	$R^2=.78, .94$	4.78(s)
Chen and Lux (2015)	3	d:1/1980–12/2010	s/fx/g	$p-v \in \langle 4.6\%, 45.5\% \rangle$	-
Chiarella et al. (2014)	6	m:502,251 o.	s	-	.44(s),.54(s),.69(s)
Chiarella et al. (2015)	5	w:2007–4/2013	CDS spreads	-	.74(i)–6.84(s)
de Jong et al. (2009a)	10	q:112 o.	s	-	1.03(s),2.87(s)
ter Ellen and Zwinkels (2010)	7	m:295,319	c (crude oil)	-	1.19(s),1.36(s)
ter Ellen et al. (2013)	2–5	w:1/2003–2/2008	fx	$adjR^2$ up to .7	7.72(i)–454.4(i)
Franke (2009)	6	d:4115–6867 o.	s,fx	$p-v \in \langle 0\%, 2\% \rangle$	-
Frijns et al. (2010)	5	d:01–12/2000	s (index options)	-	107.34(i)
Franke and Westerhoff (2011)	6	d:6866,6861 o.	s,fx	$p-v=12.8\%, 27.7\%$	-
Franke and Westerhoff (2012)	9	d:6866 o.	s	$p-v=12.7\%–32.6\%$	-
Ghoshadze and Lux (2015)	3	d:1/1980–12/2009	s/fx/g	$p-v \in \langle .3\%, 67\% \rangle$	-
Grazzini et al. (2013)	3	-	-	-	-
Grazzini and Richiardi (2015)	1	d:400 o.	s	-	-
Goldbaum and Zwinkels (2014)	4	m:2825–2941 o.	fx (experts' forecasts)	$adjR^2=.55-.79$	-
Hommes and Veld (2015)	4	q:252 o.	re	$R^2=.95$	2.44(i)
Huisman et al. (2010)	4	d:694,753,1038 o.	c (electricity futures)	-	1.06(s),1.77(s),15.87(i)
Kouwenberg and Zwinkels (2014)	4	q:127,198 o.	re	-	2.98(s),1.36(s)
Kouwenberg and Zwinkels (2015)	5	q:204 o.	re	-	2.18(s)
Lof (2012)	7	q:208 o.	s	$R^2=.97$	7.45(s), 4.74(s)
Lof (0)	5	a:140 o.	s	$R^2=.55$	type-specific: .8(i),1.13(s),5.18(i)
Reitz and Slopek (2009)	6	m:252 o.	c (crude oil)	-	-
Recchioni et al. (2015)	4	d:245 o.	s	-	2.14(s),.59(i),.03(s),.36(i)
Verschoor and Zwinkels (2013)	5	m:107 o.	fx	-	2.64(i),14.51(i)

*Note:* The Table complements information in Table 2 following the logic of Table 1. Authors are alphabetised. ‘#’ displays total number of estimated parameters; ‘Data’ describes data frequency: ‘d/w/m/q/a’ for daily/weekly/monthly/quarterly/annual, and number of observations (when a specific figure is not provided, we report starting and final years); ‘Type’ shows the type of data: ‘s/fx/c/g/re’ for stock markets/FX/commodity markets/gold/real estate; ‘Fit’ reports the statistical fit of the estimation ( $R^2$ , its alternatives, p-value of the J-test of overidentifying restrictions to accept the model as a possible data generating process); and ‘|IOC|’ displays the absolute estimated value of the ‘intensity of choice’—the switching parameter from the multinomial logit model, see Equation 16 (where relevant), furthermore ‘s’/‘i’ denotes its statistical significance/insignificance at 5% level. Figures are rounded to 2 decimal digits.

Table 10: Results for  $\beta$  estimation with normal noise

$\beta$	(a) $\hat{\beta}, N(0, 10^{-16})$						(b) $\hat{\beta}, N(0, 10^{-14})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	.00	.03	-.02	.04	41%	-.00	-.00	.07	-.15	.11	2%
.1	.10	.10	.02	.08	.12	17%	.10	.10	.05	-.03	.21	0%
.5	.50	.50	.05	.47	.54	54%	.50	.50	.11	.40	.67	7%
1	1.00	.99	.08	.91	1.05	70%	1.00	1.00	.18	.84	1.15	22%
3	3.00	3.02	.17	2.88	3.38	86%	3.00	3.01	.35	2.77	3.35	43%
5	5.00	5.08	.90	4.88	5.20	89%	5.00	4.99	.29	4.74	5.16	59%
10	10.00	9.98	.07	9.78	10.10	95%	10.00	9.99	.29	9.57	10.54	72%
$\beta$	(c) $\hat{\beta}, N(0, 10^{-12})$						(d) $\hat{\beta}, N(0, 10^{-10})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	-.01	.17	-.46	.41	0%	.00	.01	.24	-.49	.49	0%
.1	.10	.10	.11	-.10	.30	0%	.09	.09	.13	-.10	.30	0%
.5	.50	.49	.26	-.19	1.15	0%	.50	.49	.36	-.37	1.29	0%
1	1.00	.99	.33	.28	1.78	0%	1.00	1.00	.50	-.22	2.25	0%
3	3.00	3.01	.60	2.21	3.78	3%	3.01	3.02	1.07	.39	5.31	0%
5	5.00	4.97	.62	4.12	5.59	10%	5.01	4.97	1.39	2.51	6.83	2%
10	10.00	9.96	1.04	8.91	11.00	33%	10.01	9.80	2.10	5.93	11.29	20%
$\beta$	(e) $\hat{\beta}, N(0, 10^{-8})$						(f) $\hat{\beta}, N(0, 10^{-6})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.23	-.48	.48	0%	-.00	-.01	.23	-.46	.46	0%
.1	.11	.11	.12	-.10	.30	0%	.10	.10	.12	-.10	.30	0%
.5	.50	.49	.33	-.30	1.21	0%	.49	.47	.35	-.37	1.26	0%
1	1.01	1.04	.50	-.09	2.36	0%	1.01	1.04	.51	-.10	2.43	0%
3	3.01	3.03	.91	1.33	5.07	0%	2.99	3.00	.95	.93	5.03	0%
5	4.99	5.01	1.27	3.13	6.90	2%	5.00	4.98	1.19	2.52	6.52	2%
10	10.00	10.02	2.28	7.85	12.48	19%	9.99	9.94	2.00	6.83	11.61	19%
$\beta$	(g) $\hat{\beta}, N(0, 0.01^2)$						(h) $\hat{\beta}, N(0, 0.1^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.01	-.00	.23	-.49	.45	0%	.01	.01	.22	-.45	.46	0%
.1	.09	.10	.12	-.10	.30	0%	.11	.11	.12	-.10	.30	0%
.5	.50	.49	.34	-.32	1.26	0%	.50	.50	.35	-.35	1.30	0%
1	.99	.99	.52	-.35	2.31	0%	.99	1.00	.50	-.19	2.46	0%
3	3.01	3.00	.89	1.04	4.74	0%	2.99	3.05	1.00	1.48	5.86	0%
5	5.01	5.01	1.26	2.39	7.11	2%	4.99	5.05	1.21	3.75	6.81	1%
10	10.00	9.85	2.42	5.97	11.90	13%	9.99	9.99	2.22	7.57	11.64	3%
$\beta$	(i) $\hat{\beta}, N(0, 1)$						(j) $\hat{\beta}, N(0, 2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.11	-.24	.23	0%	-.00	-.00	.05	-.08	.08	0%
.1	.11	.11	.08	-.09	.30	0%	.10	.10	.04	.01	.20	0%
.5	.50	.51	.14	.23	.81	0%	.50	.51	.11	.33	.72	2%
1	1.00	1.01	.23	.66	1.45	1%	1.01	1.05	.27	.71	1.76	4%
3	3.07	3.59	1.41	2.35	7.93	3%	3.34	4.01	1.69	2.14	8.49	35%
5	5.61	7.23	3.30	3.82	14.41	8%	4.96	5.01	1.64	2.57	8.44	64%
10	11.20	13.43	6.31	5.16	28.13	23%	7.77	5.63	5.87	-9.53	10.64	96%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distributions of given parameters,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 11: Results for  $\beta$  estimation with normal noise,  $R = 1.001$ 

$\beta$	(a) $\hat{\beta}, N(0, 10^{-16})$						(b) $\hat{\beta}, N(0, 10^{-14})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.04	-.00	.01	83%	-.00	-.00	.02	-.03	.02	39%
.1	.10	.10	.01	.10	.11	72%	.10	.10	.02	.08	.12	17%
.5	.50	.50	.01	.49	.52	90%	.50	.50	.05	.47	.55	56%
1	1.00	1.01	.10	.97	1.03	93%	1.00	1.00	.10	.95	1.11	69%
3	3.00	3.00	.03	2.94	3.15	96%	3.00	3.00	.03	2.94	3.07	86%
5	5.00	5.03	.23	4.49	5.84	98%	5.00	5.05	.44	4.86	5.28	90%
10	10.00	10.01	.01	10.00	10.04	99%	10.00	9.65	2.50	9.07	10.08	94%
$\beta$	(c) $\hat{\beta}, N(0, 10^{-12})$						(d) $\hat{\beta}, N(0, 10^{-10})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.08	-.12	.16	2%	.00	.00	.18	-.46	.44	0%
.1	.10	.10	.05	-.07	.24	0%	.10	.10	.10	-.10	.30	0%
.5	.50	.50	.11	.34	.66	7%	.50	.49	.24	-.13	1.01	0%
1	1.00	.99	.15	.85	1.17	19%	1.00	1.00	.28	.49	1.58	0%
3	3.00	2.99	.29	2.77	3.21	43%	3.00	2.98	.41	2.18	3.60	3%
5	5.00	5.01	.21	4.74	5.38	59%	5.00	4.98	.84	4.19	5.76	10%
10	10.00	9.97	.47	9.60	10.24	73%	10.00	9.96	.94	9.08	10.88	30%
$\beta$	(e) $\hat{\beta}, N(0, 10^{-8})$						(f) $\hat{\beta}, N(0, 10^{-6})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.25	-.50	.50	0%	.00	.00	.23	-.46	.48	0%
.1	.09	.09	.13	-.10	.30	0%	.10	.10	.12	-.10	.30	0%
.5	.50	.50	.34	-.25	1.32	0%	.50	.51	.34	-.29	1.35	0%
1	.98	.97	.52	-.32	2.18	0%	1.02	1.03	.52	-.37	2.41	0%
3	3.00	2.97	.87	.79	4.40	0%	3.00	2.98	1.01	.44	4.86	0%
5	5.00	4.97	1.22	3.10	6.74	3%	4.99	4.96	1.56	1.59	7.41	3%
10	9.99	9.95	1.55	7.98	11.45	19%	9.99	10.04	2.04	8.29	11.97	17%
$\beta$	(g) $\hat{\beta}, N(0, 0.01^2)$						(h) $\hat{\beta}, N(0, 0.1^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.23	-.47	.45	0%	-.01	-.01	.22	-.47	.43	0%
.1	.10	.10	.12	-.10	.30	0%	.10	.10	.12	-.10	.30	0%
.5	.50	.50	.35	-.30	1.30	0%	.50	.50	.36	-.32	1.35	0%
1	1.00	.99	.55	-.43	2.40	0%	1.01	1.02	.52	-.28	2.38	0%
3	3.01	3.04	1.02	.87	5.48	0%	3.00	3.00	.90	1.16	4.85	0%
5	5.00	4.98	1.51	2.77	7.55	3%	4.99	4.98	1.41	2.90	7.09	0%
10	10.01	9.87	2.31	4.97	11.58	16%	10.01	10.07	1.83	8.82	12.20	1%
$\beta$	(i) $\hat{\beta}, N(0, 1)$						(j) $\hat{\beta}, N(0, 2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.10	-.23	.22	0%	.00	.00	.05	-.08	.10	0%
.1	.10	.10	.09	-.10	.30	0%	.10	.10	.04	-.00	.18	0%
.5	.50	.50	.14	.24	.80	0%	.50	.51	.10	.36	.73	2%
1	1.00	1.02	.20	.66	1.43	1%	1.01	1.06	.31	.71	2.03	4%
3	3.08	3.58	1.38	2.36	7.79	4%	3.40	3.93	1.57	2.11	8.36	34%
5	5.54	7.21	3.25	3.77	14.39	6%	5.03	5.25	1.67	2.80	9.13	63%
10	10.87	12.84	6.09	4.87	27.55	25%	6.14	3.97	6.82	-9.84	11.22	97%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distributions of given parameters,  $R = 1.001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 12: Results for  $\beta$  estimation with normal noise, off-centered

$\beta$	(a) $\hat{\beta}$ , $N(0, 10^{-16})$						(b) $\hat{\beta}$ , $N(0, 10^{-14})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.03	-.02	.02	40%	-.00	.00	.07	-.12	.12	2%
.1	.10	.10	.02	.09	.11	18%	.10	.10	.05	-.04	.21	0%
.5	.50	.50	.09	.46	.54	57%	.50	.51	.12	.36	.69	9%
1	1.00	1.00	.04	.95	1.04	72%	1.00	.99	.13	.80	1.13	22%
3	3.00	3.00	.31	2.89	3.07	85%	3.00	2.99	.31	2.78	3.25	45%
5	5.00	5.00	.06	4.88	5.19	90%	5.00	5.00	.60	4.56	5.30	59%
10	10.00	10.07	.49	9.93	10.17	94%	10.00	10.01	.26	9.44	10.57	74%
$\beta$	(c) $\hat{\beta}$ , $N(0, 10^{-12})$						(d) $\hat{\beta}$ , $N(0, 10^{-10})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	.01	.17	-.38	.52	0%	.03	.05	.25	-.38	.60	0%
.1	.10	.11	.10	-.05	.35	0%	.13	.13	.13	-.05	.35	0%
.5	.50	.52	.26	-.02	1.34	0%	.50	.53	.36	-.20	1.49	0%
1	1.00	1.01	.31	.37	1.79	0%	1.00	1.05	.55	-.08	2.69	0%
3	3.00	3.01	.58	2.26	3.85	4%	3.00	3.05	1.00	1.25	5.50	0%
5	5.00	5.02	.87	4.13	5.81	12%	5.00	5.08	1.39	3.08	9.45	2%
10	10.00	10.09	1.53	9.01	11.66	34%	10.00	10.16	2.47	7.95	13.38	20
$\beta$	(e) $\hat{\beta}$ , $N(0, 10^{-8})$						(f) $\hat{\beta}$ , $N(0, 10^{-6})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.01	.23	-.38	.54	0%	.03	.06	.24	-.37	.59	0%
.1	.11	.12	.12	-.05	.35	0%	.12	.13	.12	-.05	.35	0%
.5	.50	.53	.34	-.15	1.47	0%	.51	.56	.36	-.14	1.55	0%
1	1.01	1.03	.49	-.02	2.34	0%	1.01	1.07	.51	.03	2.52	0%
3	3.01	3.07	.97	1.21	5.96	0%	3.00	3.08	1.01	1.14	5.88	0%
5	5.00	5.05	1.37	2.68	7.51	3%	4.99	5.08	1.28	3.51	7.24	2%
10	10.01	10.18	2.38	7.95	14.78	20%	10.00	10.24	2.60	8.40	13.99	20%
$\beta$	(g) $\hat{\beta}$ , $N(0, 0.01^2)$						(h) $\hat{\beta}$ , $N(0, 0.1^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.02	.03	.22	-.38	.52	0%	.01	.03	.23	-.36	.56	0%
.1	.11	.13	.12	-.05	.35	0%	.12	.13	.12	-.05	.35	0%
.5	.51	.55	.35	-.13	1.50	0%	.50	.54	.35	-.15	1.54	0%
1	.99	1.05	.53	-.05	2.60	0%	1.01	1.07	.55	-.06	2.81	0%
3	3.00	3.13	1.08	1.75	6.87	0%	3.00	3.07	1.05	1.20	5.77	0%
5	5.00	5.12	1.55	2.96	8.68	3%	5.01	5.19	1.58	3.89	9.97	0%
10	9.99	9.93	2.46	5.26	12.04	17%	10.00	10.09	1.75	8.70	11.34	1%
$\beta$	(i) $\hat{\beta}$ , $N(0, 1)$						(j) $\hat{\beta}$ , $N(0, 2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.10	-.21	.20	0%	.00	.00	.05	-.08	.10	0%
.1	.10	.11	.08	-.05	.32	0%	.10	.10	.04	.02	.20	0%
.5	.50	.51	.15	.23	.89	0%	.50	.51	.10	.34	.71	2%
1	1.01	1.03	.25	.73	1.41	0%	1.01	1.09	.40	.71	2.70	5%
3	3.17	4.17	2.08	2.39	9.67	3%	3.54	4.26	1.89	2.07	9.01	40%
5	6.39	8.36	4.08	4.03	16.81	10%	5.01	5.28	1.70	2.87	9.55	67%
10	11.82	14.99	7.33	6.69	32.74	33%	6.30	6.46	3.08	-4.79	11.83	97%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distributions of given parameters,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 13: Results for  $\beta$  estimation with uniform noise I...

... covering <i>equal intervals</i> as their respective normal counterparts by 99.74% of the probability mass												
$\beta$	(a) $\widehat{\beta}$ , $U(-3 \times 10^{-8}, 3 \times 10^{-8})$						(b) $\widehat{\beta}$ , $U(-3 \times 10^{-7}, 3 \times 10^{-7})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.04	-.03	.03	26%	.00	.00	.07	-.12	.13	0%
.1	.10	.10	.02	.09	.12	10%	.10	.10	.05	-.01	.24	0%
.5	.50	.50	.04	.47	.53	42%	.50	.50	.11	.30	.73	2%
1	1.00	1.00	.09	.94	1.04	56%	1.00	1.00	.10	.86	1.12	9%
3	3.00	3.01	.11	2.96	3.14	80%	3.00	3.01	.16	2.81	3.15	31%
5	5.00	5.00	.05	4.91	5.12	85%	5.00	5.00	.39	4.85	5.22	44%
10	10.00	10.02	.20	9.88	10.07	92%	10.00	10.02	.35	9.79	10.41	62%
(c) $\widehat{\beta}$ , $U(-3 \times 10^{-6}, 3 \times 10^{-6})$							(d) $\widehat{\beta}$ , $U(-3 \times 10^{-5}, 3 \times 10^{-5})$					
0	-.00	-.01	.15	-.41	.34	0%	-.00	-.01	.20	-.45	.45	0%
.1	.10	.10	.10	-.10	.30	0%	.10	.10	.11	-.10	.30	0%
.5	.50	.50	.21	.02	.98	0%	.49	.48	.27	-.21	1.14	0%
1	1.00	1.01	.28	.49	1.59	0%	1.01	1.00	.38	-.08	1.92	0%
3	3.00	3.02	.50	2.57	3.74	1%	2.99	2.98	.79	1.76	4.19	0%
5	5.00	5.03	.60	4.44	5.69	5%	5.00	4.96	.86	4.02	5.64	2%
10	10.00	10.02	.71	9.50	10.91	28%	10.00	10.05	1.25	9.31	10.79	22%
(e) $\widehat{\beta}$ , $U(-3 \times 10^{-4}, 3 \times 10^{-4})$							(f) $\widehat{\beta}$ , $U(-3 \times 10^{-3}, 3 \times 10^{-3})$					
0	.00	-.00	.17	-.39	.39	0%	.00	.00	.17	-.39	.41	0%
.1	.10	.10	.10	-.09	.28	0%	.10	.10	.09	-.09	.29	0%
.5	.50	.49	.26	-.20	1.16	0%	.50	.49	.29	-.33	1.16	0%
1	1.00	1.00	.40	-.06	2.03	0%	1.01	1.03	.35	.29	1.96	0%
3	2.99	3.00	.73	2.13	3.69	0%	3.00	3.02	.80	2.01	4.38	0%
5	5.00	4.95	1.07	3.23	6.01	2%	4.99	5.00	1.11	4.14	5.84	2%
10	10.00	10.02	1.55	9.38	11.20	19%	10.00	9.91	1.72	8.76	10.67	17%
(g) $\widehat{\beta}$ , $U(-0.03, 0.03)$							(h) $\widehat{\beta}$ , $U(-0.3, 0.3)$					
0	-.00	-.01	.17	-.41	.37	0%	.00	-.00	.16	-.40	.35	0%
.1	.10	.10	.10	-.08	.29	0%	.10	.10	.09	-.09	.29	0%
.5	.50	.50	.26	-.21	1.11	0%	.50	.51	.25	-.15	1.14	0%
1	1.00	.99	.40	-.21	1.88	0%	1.00	.99	.40	-.08	2.17	0%
3	3.00	3.03	.72	1.97	4.56	0%	2.99	3.00	.64	2.28	3.73	0%
5	5.00	5.02	1.13	3.85	6.25	2%	5.00	5.03	.89	4.44	5.81	0%
10	10.00	10.04	1.65	9.34	10.64	12%	10.01	10.10	1.18	9.40	10.89	1%
(i) $\widehat{\beta}$ , $U(-3, 3)$							(j) $\widehat{\beta}$ , $U(-6, 6)$					
0	-.00	.00	.04	-.06	.06	0%	-.00	.00	.01	-.02	.02	0%
.1	.10	.10	.03	.03	.16	0%	.10	.10	.01	.08	.13	0%
.5	.50	.50	.05	.42	.59	1%	.50	.50	.06	.43	.61	7%
1	1.00	1.01	.15	.86	1.20	3%	1.01	1.06	.27	.81	1.99	20%
3	3.09	3.95	1.77	2.56	8.49	11%	3.14	3.40	.99	2.03	5.90	67%
5	5.56	7.28	3.17	3.95	14.41	26%	4.24	4.05	2.02	-3.68	7.20	88%
10	10.26	11.46	4.78	5.44	24.34	56%	-	-	-	-	-	100%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from uniform distributions of given parameters,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 14: Results for  $\beta$  estimation with uniform noise II...

...with equal variances as their respective normal counterparts												
$\beta$	(a) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-8}, \frac{\sqrt{12}}{2} \times 10^{-8})$						(b) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-7}, \frac{\sqrt{12}}{2} \times 10^{-7})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.02	-.01	.01	39%	-.00	-.00	.05	-.09	.10	1%
.1	.10	.10	.01	.09	.11	18%	.10	.10	.03	.04	.18	0%
.5	.50	.50	.02	.48	.53	57%	.50	.50	.08	.43	.59	6%
1	1.00	1.00	.04	.97	1.04	68%	1.00	1.00	.10	.92	1.11	20%
3	3.00	2.93	.54	2.93	3.06	87%	3.00	3.00	.15	2.84	3.12	45%
5	5.00	5.01	.07	4.95	5.11	88%	5.00	4.99	.18	4.72	5.18	55%
10	10.00	10.00	.05	9.89	10.18	94%	10.00	10.01	.17	9.85	10.29	72%
$\beta$	(c) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-6}, \frac{\sqrt{12}}{2} \times 10^{-6})$						(d) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-5}, \frac{\sqrt{12}}{2} \times 10^{-5})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	.00	.13	-.29	.39	0%	-.00	-.00	.20	-.48	.44	0%
.1	.10	.10	.09	-.10	.30	0%	.10	.10	.11	-.10	.30	0%
.5	.50	.49	.19	.00	.89	0%	.50	.51	.27	-.16	1.25	0%
1	1.00	1.00	.25	.45	1.43	0%	1.00	.98	.41	-.12	2.01	0%
3	3.00	3.00	.35	2.62	3.42	3%	3.00	3.01	.74	2.15	4.04	0%
5	5.00	4.97	.54	4.44	5.41	12%	5.01	5.00	.86	3.84	5.84	2%
10	10.00	10.01	1.04	9.49	10.46	32%	10.00	10.02	1.38	9.23	10.90	21%
$\beta$	(e) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-4}, \frac{\sqrt{12}}{2} \times 10^{-4})$						(f) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-3}, \frac{\sqrt{12}}{2} \times 10^{-3})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.18	-.40	.42	0%	.00	.00	.17	-.39	.42	0%
.1	.10	.10	.10	-.09	.29	0%	.10	.10	.09	-.09	.29	0%
.5	.50	.50	.28	-.25	1.22	0%	.50	.50	.27	-.23	1.25	0%
1	1.00	1.02	.44	-.03	2.36	0%	1.00	1.00	.39	-.06	2.08	0%
3	3.00	3.00	.80	1.48	4.29	0%	3.00	2.99	.74	2.01	3.85	0%
5	5.00	4.95	1.03	3.79	5.72	3%	5.00	5.03	.87	4.25	6.43	2%
10	10.01	9.93	1.79	9.41	10.74	19%	10.00	9.97	1.76	9.11	10.69	19%
$\beta$	(g) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-2}, \frac{\sqrt{12}}{2} \times 10^{-2})$						(h) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.17	-.40	.39	0%	-.00	-.00	.17	-.40	.39	0%
.1	.10	.10	.09	-.08	.29	0%	.10	.10	.10	-.09	.29	0%
.5	.50	.52	.27	-.15	1.21	0%	.50	.49	.27	-.21	1.19	0%
1	1.00	.99	.45	-.23	2.21	0%	1.00	.99	.36	.13	1.91	0%
3	3.00	3.01	.72	2.15	4.44	0%	3.01	3.02	.80	1.71	4.55	0%
5	5.00	4.98	.90	4.08	5.87	3%	5.00	5.05	.97	4.51	5.81	1%
10	9.99	9.99	1.34	9.27	10.74	14%	10.01	10.02	1.51	9.44	10.71	2%
$\beta$	(i) $\hat{\beta}$ , $U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2})$						(j) $\hat{\beta}$ , $U(-2\frac{\sqrt{12}}{2}, 2\frac{\sqrt{12}}{2})$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.06	-.11	.15	0%	.00	.00	.02	-.04	.04	0%
.1	.10	.10	.06	-.02	.25	0%	.10	.10	.03	.06	.16	0%
.5	.50	.50	.09	.34	.64	0%	.50	.50	.06	.42	.59	2%
1	1.00	1.00	.08	.82	1.16	1%	1.00	1.01	.11	.87	1.18	3%
3	3.02	3.18	.82	2.64	6.53	3%	3.16	4.25	1.96	2.54	8.67	18%
5	5.13	6.62	2.89	4.34	13.75	6%	5.49	6.81	2.86	3.49	13.82	38%
10	11.05	14.78	6.59	8.05	29.42	17%	9.46	9.02	4.31	-5.54	17.48	68%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from uniform distributions of given parameters,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.



Table 15: Results for  $\beta$  estimation w.r.t. various dist. of  $g_h$  &  $b_h$  I.

$\beta$	(a) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 0.1^2)$						(b) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 0.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.02	.01	.29	-.47	.48	0%	-.02	-.01	.28	-.48	.48	0%
.1	.10	.10	.12	-.09	.29	0%	.10	.10	.12	-.10	.30	0%
.5	.50	.51	.57	-.42	1.44	0%	.51	.51	.48	-.40	1.39	0%
1	.87	.94	1.15	-.95	2.91	0%	.99	.99	.80	-.73	2.73	0%
3	2.95	2.95	3.25	-2.67	8.69	0%	3.00	2.97	1.81	-1.65	7.43	0%
5	4.99	5.07	5.04	-4.26	14.17	0%	4.97	4.77	2.60	-3.23	10.94	0%
10	10.02	10.15	9.03	-8.22	27.82	0%	9.98	9.92	3.65	.31	18.84	0%
$\beta$	(c) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 0.3^2)$						(d) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 0.4^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.23	-.47	.47	0%	-.01	-.00	.17	-.38	.43	0%
.1	.10	.10	.12	-.10	.30	0%	.10	.10	.11	-.10	.30	0%
.5	.50	.50	.34	-.34	1.27	0%	.50	.50	.24	-.09	1.17	0%
1	1.01	1.01	.54	-.40	2.33	0%	1.00	1.00	.37	.20	1.85	0%
3	3.00	3.01	.83	1.57	4.93	0%	2.99	2.95	.78	1.24	3.76	0%
5	5.00	5.00	1.35	3.02	6.84	0%	5.00	5.00	.85	4.17	5.65	1%
10	9.99	10.01	1.87	8.19	11.55	1%	10.00	10.02	1.33	9.11	10.77	6%
$\beta$	(e) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 0.6^2)$						(f) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 0.8^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	-.00	.11	-.34	.25	0%	.00	.00	.07	-.14	.16	0%
.1	.10	.10	.08	-.09	.29	0%	.10	.10	.05	-.00	.25	1%
.5	.50	.50	.13	.28	.75	0%	.50	.50	.08	.39	.62	2%
1	1.00	1.00	.15	.71	1.27	0%	1.00	1.00	.11	.84	1.16	6%
3	3.00	3.00	.19	2.72	3.29	7%	3.00	3.01	.27	2.84	3.18	29%
5	5.00	4.99	.31	4.62	5.26	16%	5.00	5.01	.29	4.80	5.27	42%
10	10.01	10.04	.88	9.55	10.36	28%	10.00	10.04	1.05	9.47	10.28	54%
$\beta$	(g) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 1)$						(h) $\hat{\beta}$ , $g_h$ & $b_h \sim N(0, 1.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.05	-.06	.10	2%	.00	.00	.03	-.05	.06	7%
.1	.10	.10	.04	.02	.18	2%	.10	.10	.03	.05	.15	4%
.5	.50	.50	.07	.39	.58	9%	.50	.50	.04	.43	.56	20%
1	1.00	1.00	.09	.91	1.13	20%	1.00	1.00	.05	.94	1.09	40%
3	3.00	3.00	.07	2.84	3.13	52%	3.00	2.99	.14	2.85	3.09	69%
5	5.00	5.01	.26	4.87	5.20	64%	5.00	5.04	.67	4.59	5.18	75%
10	10.00	10.21	1.78	9.59	11.98	73%	10.01	10.42	2.59	9.72	21.17	83%

*Note:* Belief parameters  $g_h$  and  $b_h$  drawn from various normal distributions of given parameter, stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0, 0.1^2)$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 16: Results for  $\beta$  estimation w.r.t. various dist. of  $g_h$  &  $b_h$  II.

$\beta$	(a) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 0.1^2)$						(b) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 0.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.02	.01	.29	-.47	.48	0%	-.00	-.00	.24	-.47	.49	0%
.1	.10	.10	.12	-.09	.29	0%	.10	.10	.13	-.10	.30	0%
.5	.51	.52	.57	-.44	1.46	0%	.51	.53	.37	-.34	1.33	0%
1	1.02	1.05	1.11	-.85	2.92	0%	1.02	1.08	.58	-.12	2.64	0%
3	3.62	3.55	3.22	-2.50	8.75	0%	3.49	4.28	2.17	.83	8.59	0%
5	6.74	6.39	5.62	-4.34	14.58	0%	6.59	7.86	3.89	2.14	14.57	0%
10	12.73	12.32	11.46	-8.31	29.60	1%	12.51	13.98	8.89	-5.54	29.58	2%
	(c) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 0.3^2)$						(d) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 0.4^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.16	-.39	.34	0%	-.00	-.00	.09	-.21	.19	0%
.1	.10	.10	.10	-.10	.30	0%	.10	.10	.07	-.06	.26	0%
.5	.51	.52	.19	.17	.97	0%	.50	.50	.11	.31	.73	0%
1	1.00	1.02	.30	.60	1.70	0%	1.00	1.01	.19	.75	1.32	1%
3	3.17	3.78	1.57	2.33	8.23	0%	3.05	3.43	1.25	2.48	7.72	4%
5	5.82	7.43	3.42	3.71	14.46	1%	5.35	6.92	3.15	3.89	14.39	8%
10	11.49	13.91	6.99	-.28	28.97	9%	11.16	13.21	5.83	6.10	28.03	25%
	(e) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 0.6^2)$						(f) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 0.8^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.04	-.08	.08	0%	.00	-.00	.02	-.05	.04	0%
.1	.10	.10	.04	.02	.20	0%	.10	.10	.02	.05	.14	7%
.5	.50	.50	.06	.39	.63	12%	.50	.50	.04	.42	.59	30%
1	1.00	1.00	.09	.83	1.23	16%	1.00	1.00	.07	.87	1.15	41%
3	3.02	3.17	.78	2.56	6.19	26%	3.02	3.08	.53	2.57	3.64	62%
5	5.29	6.67	2.84	4.20	13.73	42%	5.10	6.14	2.51	4.16	13.49	64%
10	10.52	12.26	4.88	6.59	25.70	62%	10.29	11.94	4.29	7.88	24.38	84%
	(g) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 1)$						(h) $\hat{\beta}, g_h \text{ \& } b_h \sim N(0, 1.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	.00	.01	-.03	.03	3%	.00	.00	.01	-.02	.02	9%
.1	.10	.10	.01	.07	.13	20%	.10	.10	.01	.08	.13	38%
.5	.50	.50	.03	.45	.58	55%	.50	.50	.02	.45	.54	68%
1	1.00	1.00	.06	.89	1.10	64%	1.00	1.00	.04	.91	1.09	77%
3	2.99	3.00	.30	2.61	3.34	75%	2.98	3.01	.31	2.65	3.41	87%
5	5.14	6.19	2.56	4.24	13.26	82%	5.04	6.17	2.69	4.36	13.96	93%
10	10.26	11.42	3.19	8.08	20.32	92%	9.85	10.75	2.88	8.47	19.13	96%

*Note:* Belief parameters  $g_h$  and  $b_h$  drawn from various normal distributions of given parameter, stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0, 1)$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 17: Results for  $\beta$  estim. w.r.t. various dist. of  $g_h$  and  $b_h$  III.

$\beta$	(a) $\hat{\beta}, g_h \& b_h \sim N(0, 0.1^2)$						(b) $\hat{\beta}, g_h \& b_h \sim N(0, 0.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.00	.00	.27	-.50	.50	0%	.00	.00	.13	-.36	.39	0%
.1	.10	.10	.13	-.10	.30	0%	.10	.10	.08	-.10	.30	0%
.5	.50	.49	.42	-.50	1.46	0%	.50	.50	.17	.13	.83	0%
1	.99	.98	.63	-.62	2.70	0%	1.00	1.00	.24	.58	1.37	3%
3	3.02	3.08	1.24	-.20	6.18	0%	3.00	3.03	.38	2.65	3.59	15%
5	4.99	4.89	1.67	.58	8.08	1%	5.00	4.96	.88	4.27	5.51	28%
10	10.00	9.96	2.85	3.87	14.45	4%	10.00	9.89	1.26	8.84	10.74	46%
$\beta$	(c) $\hat{\beta}, g_h \& b_h \sim N(0, 0.3^2)$						(d) $\hat{\beta}, g_h \& b_h \sim N(0, 0.4^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.07	-.15	.11	1%	.00	.00	.04	-.06	.06	10%
.1	.10	.10	.05	-.05	.23	0%	.10	.10	.03	.04	.17	2%
.5	.50	.50	.11	.35	.68	7%	.50	.50	.06	.43	.57	24%
1	1.00	.99	.17	.77	1.19	18%	1.00	1.00	.13	.84	1.08	40%
3	3.00	3.00	.12	2.79	3.18	46%	3.00	2.99	.33	2.79	3.21	64%
5	5.00	5.03	.39	4.77	5.32	56%	5.00	5.01	.13	4.82	5.20	74%
10	10.00	9.94	1.21	9.63	10.29	74%	10.00	10.00	.29	9.42	10.31	85%
$\beta$	(e) $\hat{\beta}, g_h \& b_h \sim N(0, 0.6^2)$						(f) $\hat{\beta}, g_h \& b_h \sim N(0, 0.8^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.02	-.02	.02	36%	-.00	-.00	.02	-.03	.02	54%
.1	.10	.10	.01	.08	.12	15%	.10	.10	.01	.08	.11	35%
.5	.50	.50	.05	.46	.55	51%	.50	.50	.02	.47	.52	68%
1	1.00	1.00	.05	.92	1.07	67%	1.00	1.00	.01	.98	1.02	83%
3	3.00	3.01	.09	2.93	3.09	86%	3.00	3.02	.18	2.94	3.09	94%
5	5.00	5.05	.32	4.90	5.42	91%	5.00	5.00	.03	4.92	5.07	97%
10	10.00	9.97	.60	9.76	10.38	96%	9.99	10.28	4.14	.80	25.02	98%
$\beta$	(g) $\hat{\beta}, g_h \& b_h \sim N(0, 1)$						(h) $\hat{\beta}, g_h \& b_h \sim N(0, 1.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.01	-.02	.01	72%	.00	.00	.03	-.01	.01	79%
.1	.10	.10	.01	.09	.11	53%	.10	.10	.00	.10	.10	67%
.5	.50	.50	.01	.48	.51	78%	.50	.50	.02	.47	.51	88%
1	1.00	1.00	.03	.97	1.02	89%	1.00	1.00	.02	.98	1.11	94%
3	3.00	3.30	1.16	2.96	8.29	97%	3.00	3.16	.49	2.98	4.82	98%
5	5.00	5.00	.05	4.91	5.15	99%	5.00	5.59	1.68	4.93	9.76	99%
10	10.02	11.61	4.69	5.79	20.28	99%	10.01	14.29	7.09	9.62	29.62	99%

*Note:* Belief parameters  $g_h$  and  $b_h$  drawn from various normal distributions of given parameter, stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0, 10^{-14})$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 18: Results for  $\beta$  estim. w.r.t. various dist. of  $g_h$  and  $b_h$  IV.

$\beta$	(a) $\hat{\beta}, g_h \& b_h \sim N(0, 0.1^2)$						(b) $\hat{\beta}, g_h \& b_h \sim N(0, 0.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	.01	.00	.32	-.50	.50	0%	.01	.01	.26	-.50	.50	0%
.1	.10	.10	.13	-.10	.30	0%	.09	.09	.14	-.10	.30	0%
.5	.49	.50	.60	-.50	1.50	0%	.50	.49	.39	-.43	1.39	0%
1	.99	.98	1.06	-.96	2.93	0%	.99	1.01	.61	-.31	2.56	0%
3	3.03	3.05	2.59	-2.46	8.24	0%	3.00	3.06	1.11	.73	5.72	0%
5	5.06	5.20	3.77	-3.64	13.56	0%	5.00	4.90	1.73	-.77	7.65	0%
10	9.94	9.59	6.60	-7.20	25.75	0%	10.00	9.89	2.03	6.75	11.93	4%
$\beta$	(c) $\hat{\beta}, g_h \& b_h \sim N(0, 0.3^2)$						(d) $\hat{\beta}, g_h \& b_h \sim N(0, 0.4^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.18	-.49	.46	0%	-.00	-.00	.12	-.34	.25	0%
.1	.10	.10	.11	-.10	.30	0%	.10	.10	.08	-.10	.30	0%
.5	.50	.51	.25	-.14	1.15	0%	.50	.50	.14	.19	.76	0%
1	1.00	1.00	.35	.27	1.74	0%	1.00	.99	.23	.58	1.35	3%
3	3.00	3.01	.51	2.43	3.73	4%	3.00	3.00	.47	2.53	3.45	15%
5	5.00	5.04	.59	4.34	5.83	10%	5.00	5.00	.31	4.51	5.44	29%
10	10.01	9.98	1.27	9.23	11.22	29%	10.01	9.99	.48	9.31	10.59	60%
$\beta$	(e) $\hat{\beta}, g_h \& b_h \sim N(0, 0.6^2)$						(f) $\hat{\beta}, g_h \& b_h \sim N(0, 0.8^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.07	-.11	.11	2%	-.00	.00	.05	-.06	.08	11%
.1	.10	.10	.05	-.01	.22	0%	.10	.10	.03	.05	.14	3%
.5	.50	.50	.08	.40	.61	7%	.50	.50	.04	.43	.58	23%
1	1.00	1.00	.09	.86	1.15	18%	1.00	1.00	.04	.94	1.06	41%
3	3.00	3.00	.10	2.83	3.19	51%	3.00	3.00	.09	2.75	3.16	75%
5	5.00	5.00	.19	4.71	5.33	66%	5.00	5.01	.06	4.89	5.14	87%
10	10.00	10.12	.94	9.56	10.71	86%	10.00	9.95	.34	9.38	10.45	95%
$\beta$	(g) $\hat{\beta}, g_h \& b_h \sim N(0, 1)$						(h) $\hat{\beta}, g_h \& b_h \sim N(0, 1.2^2)$					
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN
0	-.00	-.00	.04	-.03	.04	27%	-.00	-.00	.02	-.02	.02	42%
.1	.10	.10	.02	.07	.13	12%	.10	.10	.02	.08	.12	25%
.5	.50	.50	.04	.46	.55	43%	.50	.50	.02	.47	.53	58%
1	1.00	1.00	.04	.94	1.04	62%	1.00	1.00	.05	.93	1.07	73%
3	3.00	3.00	.12	2.89	3.11	86%	3.00	3.00	.05	2.87	3.13	94%
5	5.00	5.25	1.16	4.85	10.13	94%	5.00	5.59	2.29	4.72	14.47	97%
10	10.00	10.76	3.36	5.79	22.97	97%	10.01	14.99	8.23	9.65	30.00	98%

*Note:* Belief parameters  $g_h$  and  $b_h$  drawn from various normal distributions of given parameter, stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0, 10^{-12})$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. ‘NN’ column reports the percentage of runs with ‘NaN’ outcome rounded to integer numbers.

Table 19: Results for  $\beta$  estimation with various combined noises II.

$\beta$	(a) $\hat{\beta}, \epsilon_t \sim N(0, 1),$ $\{\varepsilon_i\}_{i=1}^N \sim U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2})$						(b) $\hat{\beta}, \epsilon_t \sim U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2}),$ $\{\varepsilon_i\}_{i=1}^N \sim N(0, 1)$						
	Med.	Mean	SD	LQ	HQ	NN	Med.	Mean	SD	LQ	HQ	NN	
0	.00	-.01	.14	-.38	.29	0%	.00	.01	.12	-.25	.24	0%	
.1	.10	.10	.10	-.10	.30	0%	.10	.10	.10	-.10	.30	0%	
.5	.50	.51	.16	.19	.89	0%	.50	.51	.14	.22	.82	0%	
1	.99	1.00	.25	.54	1.56	1%	1.00	1.01	.23	.63	1.48	1%	
3	3.05	3.22	.95	2.14	6.41	5%	3.07	3.56	1.44	2.25	7.97	3%	
5	5.41	6.68	2.97	3.44	14.03	9%	5.55	7.27	3.28	3.79	14.34	6%	
10	11.95	13.88	5.88	6.94	27.99	24%	11.39	14.03	6.84	5.51	29.30	13%	
(c) $\hat{\beta}, \epsilon_t \sim N(0, 0.1^2),$ $\{\varepsilon_i\}_{i=1}^N \sim U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1})$							(d) $\hat{\beta}, \epsilon_t \sim U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1}),$ $\{\varepsilon_i\}_{i=1}^N \sim N(0, 0.1^2)$						
0	.02	.01	.27	-.50	.50	0%	-.00	-.00	.22	-.47	.46	0%	
.1	.10	.10	.14	-.10	.30	0%	.10	.10	.12	-.10	.30	0%	
.5	.50	.50	.40	-.43	1.44	0%	.50	.50	.33	-.27	1.27	0%	
1	1.01	1.01	.53	-.21	2.39	0%	1.00	1.03	.52	-.31	2.45	0%	
3	3.02	3.03	.86	1.46	4.70	0%	3.00	3.03	1.06	.64	6.23	0%	
5	5.01	5.01	1.18	3.05	6.76	0%	5.00	5.00	1.26	3.32	6.79	0%	
10	10.00	10.01	1.43	8.13	11.88	2%	10.00	10.08	1.67	8.83	11.70	1%	
(e) $\hat{\beta}, \epsilon_t \sim N(0, 0.1^2),$ $\{\varepsilon_i\}_{i=1}^N \sim U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2})$							(f) $\hat{\beta}, \epsilon_t \sim U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1}),$ $\{\varepsilon_i\}_{i=1}^N \sim N(0, 1)$						
0	.01	.01	.28	-.47	.47	0%	-.00	-.00	.29	-.48	.48	0%	
.1	.09	.09	.12	-.09	.29	0%	.10	.09	.12	-.09	.29	0%	
.5	.46	.49	.58	-.46	1.43	0%	.46	.48	.58	-.46	1.45	0%	
1	.95	.98	1.18	-.94	2.88	0%	1.06	1.04	1.12	-.89	2.90	0%	
3	3.28	3.07	3.42	-2.75	8.68	0%	3.07	3.06	3.34	-2.73	8.81	0%	
5	4.98	5.11	5.73	-4.77	14.78	0%	5.02	5.20	5.28	-4.13	14.36	0%	
10	12.01	10.84	11.87	-10.00	29.41	1%	12.23	12.13	10.62	-8.21	29.17	1%	
(g) $\hat{\beta}, \epsilon_t \sim N(0, 1),$ $\{\varepsilon_i\}_{i=1}^N \sim U(-\frac{\sqrt{12}}{2} \times 10^{-1}, \frac{\sqrt{12}}{2} \times 10^{-1})$							(h) $\hat{\beta}, \epsilon_t \sim U(-\frac{\sqrt{12}}{2}, \frac{\sqrt{12}}{2}),$ $\{\varepsilon_i\}_{i=1}^N \sim N(0, 0.1^2)$						
						100%							100%

*Note:* Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from different distributions with various variance,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD), 2.5% (LQ), and 97.5% (HQ) quantiles are reported. Figures are rounded to 2 decimal digits. 'NN' column reports the percentage of runs with 'NaN' outcome rounded to integer numbers.

Table 20: Results of 3-parameter estimation of a 2-type model I.

True $\beta, g_2, b_2$	(a) $\hat{\beta}$			(b) $\hat{g}_2$		(c) $\hat{b}_2$		(d) $LL$			
	Md	Mn	SD	Mn	SD	Mn	SD	L-rat	$2\Delta LL$	p-v	NN
.0, .2, .15	-.01	-.00	.29	.20	.03	.15	.01	1	0	1	0%
.5, .2, .15	.49	.50	.57	.20	.03	.15	.01	1	0	1	0%
3, .2, .15	3.01	2.99	3.44	.20	.05	.15	.01	1	0	1	0%
10, .2, .15	11.60	12.19	9.46	.19	.10	.15	.01	1	0	1	0%
.0, -.2, -.15	.02	-.00	.28	-.20	.03	-.15	.01	1	0	1	0%
.5, -.2, -.15	.50	.48	.56	-.20	.03	-.15	.01	1	0	1	0%
3, -.2, -.15	3.11	2.97	3.39	-.20	.05	-.15	.01	1	0	1	0%
10, -.2, -.15	10.46	9.93	11.30	-.19	.10	-.15	.01	1	0	1	0%
.0, .2, -.15	.03	.01	.28	.20	.03	-.15	.01	1	0	1	0%
.5, .2, -.15	.56	.52	.56	.20	.03	-.15	.01	1	0	1	0%
3, .2, -.15	3.31	3.12	3.40	.20	.05	-.15	.01	1	0	1	0%
10, .2, -.15	11.97	12.23	9.60	.19	.10	-.15	.01	1	0	1	0%
.0, -.2, .15	-.00	-.01	.30	-.20	.03	.15	.01	1	0	1	0%
.5, -.2, .15	.53	.50	.60	-.20	.03	.15	.01	1	0	1	0%
3, -.2, .15	3.31	3.08	3.53	-.20	.04	.15	.01	1	0	1	0%
10, -.2, .15	11.37	10.62	11.69	-.19	.10	.15	.01	1	0	1	0%
.0, .4, .3	.00	.01	.33	.40	.03	.30	.01	1	0	1	0%
.5, .4, .3	.50	.49	.48	.40	.04	.30	.01	1	0	1	0%
3, .4, .3	2.99	3.04	.46	.40	.04	.30	.01	.99	.01	.92	0%
10, .4, .3	10.00	10.01	.61	.40	.03	.30	.01	.94	.13	.62	4%
.0, -.4, -.3	-.08	-.06	.30	-.40	.03	-.30	.01	1	0	1	0%
.5, -.4, -.3	.31	.35	.58	-.39	.04	-.30	.01	1	0	1	0%
3, -.4, -.3	2.96	2.84	1.47	-.39	.06	-.30	.01	1	0	1	0%
10, -.4, -.3	10.08	10.11	1.24	-.40	.04	-.30	.01	.99	.03	.86	2%
.0, .4, -.3	-.01	.00	.31	.40	.03	-.30	.01	1	0	1	0%
.5, .4, -.3	.51	.48	.48	.40	.04	-.30	.01	1	0	1	0%
3, .4, -.3	3.03	3.04	.46	.40	.04	-.30	.01	.99	.01	.92	1%
10, .4, -.3	9.98	10.02	.63	.40	.03	-.30	.01	.94	.13	.62	2%
.0, -.4, .3	-.06	-.05	.30	-.40	.03	.30	.01	1	0	1	0%
.5, -.4, .3	.23	.32	.58	-.40	.04	.30	.01	1	0	1	0%
3, -.4, .3	2.98	2.87	1.38	-.40	.06	.30	.01	1	0	1	0%
10, -.4, .3	9.98	10.00	1.13	-.40	.04	.30	.01	.99	.03	.86	2%
.0, .8, .6	-.00	-.00	.06	.80	.03	.60	.02	1	0	1	30%
.5, .8, .6	.50	.50	.06	.80	.03	.60	.02	.99	.02	.89	31%
3, .8, .6	2.99	3.00	.07	.80	.02	.60	.01	.59	1.07	.30	54%
10, .8, .6	9.99	9.99	.12	.80	.00	.60	.01	NA	NA	NA	95%
.0, -.8, -.6	.00	-.01	.27	-.80	.04	-.60	.01	1	0	1	17%
.5, -.8, -.6	.52	.51	.26	-.80	.04	-.60	.01	1	0	1	21%
3, -.8, -.6	3.00	3.00	.26	-.80	.04	-.60	.01	.99	.02	0	34%
10, -.8, -.6	9.98	9.99	.45	-.80	.03	-.60	.01	.89	.24	0	43%
.0, .8, -.6	.00	.00	.06	.80	.03	-.60	.02	1	0	1	32%
.5, .8, -.6	.50	.50	.06	.80	.03	-.60	.02	.99	.02	.89	33%
3, .8, -.6	3.00	3.00	.07	.80	.02	-.60	.01	.58	1.08	.30	58%
10, .8, -.6	9.97	9.99	.13	.80	.00	-.60	.00	NA	NA	NA	96%
.0, -.8, .6	-.00	-.00	.27	-.80	.04	.60	.01	1	0	1	16%
.5, -.8, .6	.52	.51	.26	-.80	.04	.60	.01	1	0	1	20%
3, -.8, .6	3.01	3.01	.27	-.80	.04	.60	.01	.99	.02	.89	31%
10, -.8, .6	10.04	10.02	.43	-.80	.03	.60	.01	.89	.24	.62	42%
.0, 1.2, .9	.00	.00	.01	1.20	.03	.90	.03	1	0	1	67%
.5, 1.2, .9	.50	.50	.01	1.20	.01	.90	.02	.70	.70	.40	77%
3, 1.2, .9											100%
10, 1.2, .9											100%
.0, -1.2, -.9	-.00	-.00	.11	-1.20	.04	-.90	.01	1	0	1	47%
.5, -1.2, -.9	.49	.50	.11	-1.20	.04	-.90	.01	1	0	1	49%
3, -1.2, -.9	2.99	3.00	.14	-1.20	.03	-.90	.01	.96	.09	.76	56%
10, -1.2, -.9	9.99	9.99	.29	-1.20	.02	-.90	.01	.60	1.03	.31	72%
.0, 1.2, -.9	.00	.00	.01	1.20	.03	-.90	.03	1	0	1	67%
.5, 1.2, -.9	.50	.50	.01	1.20	.02	-.90	.02	.70	.71	.40	76%
3, 1.2, -.9											100%
10, 1.2, -.9											100%
.0, -1.2, .9	.01	.00	.11	-1.20	.04	.90	.01	1	0	1	50%
.5, -1.2, .9	.51	.50	.11	-1.20	.04	.90	.01	1	0	1	52%
3, -1.2, .9	3.00	3.00	.13	-1.20	.04	.90	.01	.96	.09	.76	60%
10, -1.2, .9	10.00	10.00	.30	-1.20	.02	.90	.01	.60	1.03	.31	72%

Note: Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_t\}_{t=1}^N$  drawn from normal distribution  $N(0, 0.1^2)$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians (Md), means (Mn), and standard deviations (SD) are reported. Figures are rounded to 2 decimal digits. 'L-rat' denotes the likelihood ratio of the null static (i.e. restricted) model vs. the alternative switching model, ' $2\Delta LL$ ' is the test statistics of the log-likelihood ratio test being approximately  $\chi^2$  distributed with 1 degree of freedom, and 'p-v' is related p-value of the test. 'NN' column reports the percentage of runs with 'NaN' outcome rounded to integer numbers. 'NA' typically means that the static model is associated with 100% of 'NaN' outcomes.

Table 21: Results of 3-parameter estimation of a 2-type model II.

True $\beta, g_2, b_2$	(a) $\hat{\beta}$			(b) $\hat{g}_2$		(c) $\hat{b}_2$		(d) $LL$			
	Md	Mn	SD	Mn	SD	Mn	SD	L-rat	$2\Delta LL$	p-v	NN
.0, .2, .15	.03	.02	.30	.20	.03	.15	.07	1	0	1	0%
.5, .2, .15	.62	.60	.56	.19	.03	.15	.07	1	0	1	0%
3, .2, .15	3.99	4.13	2.75	.19	.03	.15	.06	1	0	1	0%
10, .2, .15	11.42	11.23	11.08	.18	.04	.14	.05	1	0	1	1%
.0, -.2, -.15	-.06	-.04	.31	-.20	.03	-.15	.07	1	0	1	0%
.5, -.2, -.15	.46	.46	.57	-.20	.03	-.15	.07	1	0	1	0%
3, -.2, -.15	3.14	3.10	3.37	-.18	.04	-.13	.07	1	0	1	0%
10, -.2, -.15	9.54	9.53	11.40	-.16	.06	-.13	.06	1	0	1	1%
.0, .2, -.15	.03	.02	.30	.20	.03	-.15	.07	1	0	1	0%
.5, .2, -.15	.62	.59	.57	.19	.03	-.15	.07	1	0	1	0%
3, .2, -.15	3.99	4.16	2.61	.19	.03	-.15	.06	1	0	1	0%
10, .2, -.15	11.33	11.43	10.87	.18	.04	-.14	.05	1	0	1	0%
.0, -.2, .15	-.05	-.04	.31	-.20	.03	.15	.07	1	0	1	0%
.5, -.2, .15	.49	.49	.56	-.20	.03	.15	.07	1	0	1	0%
3, -.2, .15	3.42	3.37	3.36	-.18	.04	.13	.06	1	0	1	0%
10, -.2, .15	10.43	10.25	11.40	-.16	.06	.12	.06	1	0	1	2%
.0, .4, .3	.00	.01	.15	.40	.03	.30	.07	1	0	1	0%
.5, .4, .3	.49	.55	.26	.40	.03	.30	.07	1	0	1	0%
3, .4, .3	3.95	4.54	2.13	.38	.04	.29	.04	.99	.02	.89	0%
10, .4, .3	12.48	13.14	9.41	.38	.05	.27	.09	.98	.04	.84	4%
.0, -.4, -.3	.00	-.01	.14	-.40	.03	-.30	.07	1	0	1	0%
.5, -.4, -.3	.49	.52	.20	-.40	.03	-.30	.07	1	0	1	0%
3, -.4, -.3	3.96	4.38	2.41	-.38	.06	-.28	.06	.99	.02	.89	0%
10, -.4, -.3	10.76	10.44	10.21	-.33	.13	-.27	.08	.99	.02	.89	11%
.0, .4, -.3	-.00	.00	.15	.40	.03	-.30	.07	1	0	1	0%
.5, .4, -.3	.50	.54	.25	.40	.03	-.30	.07	1	0	1	0%
3, .4, -.3	4.01	4.53	2.05	.38	.04	-.29	.04	.99	.02	.89	0%
10, .4, -.3	12.30	13.17	8.79	.38	.05	-.28	.08	.98	.04	.84	4%
.0, -.4, .3	-.01	-.01	.13	-.40	.03	.30	.07	1	0	1	0%
.5, -.4, .3	.49	.51	.19	-.40	.03	.30	.08	1	0	1	0%
3, -.4, .3	4.03	4.40	2.47	-.38	.06	.28	.06	.99	.02	.89	0%
10, -.4, .3	10.42	9.92	10.17	-.33	.14	.26	.09	.99	.02	.89	12%
.0, .8, .6	.00	.00	.03	.80	.03	.60	.07	1	0	1	0%
.5, .8, .6	.50	.50	.07	.80	.03	.60	.06	.97	.05	.82	3%
3, .8, .6	3.01	3.04	.37	.80	.02	.60	.03	.83	.36	.55	23%
10, .8, .6	10.24	10.54	2.16	.80	.01	.60	.02	.75	.56	.45	72%
.0, -.8, -.6	.00	.00	.02	-.80	.03	-.60	.07	1	0	1	0%
.5, -.8, -.6	.50	.51	.08	-.80	.03	-.60	.08	.98	.04	.84	0%
3, -.8, -.6	3.19	3.61	2.47	-.74	.18	-.57	.09	.94	.12	.73	4%
10, -.8, -.6	8.50	4.97	8.61	-.57	.33	-.44	.26	.95	.09	.76	44%
.0, .8, -.6	-.00	-.00	.03	.80	.03	-.60	.07	1	0	1	0%
.5, .8, -.6	.50	.50	.07	.80	.03	-.60	.06	.97	.05	.82	2%
3, .8, -.6	3.00	3.04	.36	.80	.02	-.60	.03	.83	.36	.55	23%
10, .8, -.6	10.19	10.51	2.22	.80	.01	-.60	.02	.75	.57	.45	69%
.0, -.8, .6	.00	.00	.02	-.80	.03	.60	.07	1	0	1	0%
.5, -.8, .6	.50	.50	.08	-.80	.03	.60	.07	.98	.04	.84	1%
3, -.8, .6	3.19	3.62	2.38	-.75	.17	.57	.08	.94	.12	.73	4%
10, -.8, .6	8.56	5.67	8.17	-.60	.32	.46	.25	.95	.11	.74	44%
.0, 1.2, .9	-.00	.00	.01	1.20	.02	.90	.07	1	0	1	25%
.5, 1.2, .9											100%
3, 1.2, .9											100%
10, 1.2, .9											100%
.0, -1.2, -.9	.00	.00	.01	-1.20	.03	-.90	.07	1	0	1	19%
.5, -1.2, -.9	.50	.51	.05	-1.20	.03	-.90	.07	.91	.19	.66	28%
3, -1.2, -.9	2.97	2.39	2.06	-1.07	.34	-.79	.34	.83	.38	.46	48%
10, -1.2, -.9	7.86	3.28	7.85	-.82	.50	-.54	.54	.88	.25	.62	88%
.0, 1.2, -.9	.00	.00	.01	1.20	.03	-.90	.07	1	0	1	23%
.5, 1.2, -.9											100%
3, 1.2, -.9											100%
10, 1.2, -.9											100%
.0, -1.2, .9	.00	.00	.01	-1.20	.03	.90	.07	1	0	1	18%
.5, -1.2, .9	.50	.50	.05	-1.20	.03	.90	.07	.91	.19	.66	26%
3, -1.2, .9	2.94	2.34	1.95	-1.08	.34	.81	.30	.83	.38	.46	50%
10, -1.2, .9	8.25	3.21	8.26	-.79	.52	.57	.50	.89	.23	.63	89%

Note: Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0,1)$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians (Md), means (Mn), and standard deviations (SD) are reported. Figures are rounded to 2 decimal digits. 'L-rat' denotes the likelihood ratio of the null static (i.e. restricted) model vs. the alternative switching model, ' $2\Delta LL$ ' is the test statistics of the log-likelihood ratio test being approximately  $\chi^2$  distributed with 1 degree of freedom, and 'p-v' is related p-value of the test. 'NN' column reports the percentage of runs with 'NaN' outcome rounded to integer numbers. 'NA' typically means that the static model is associated with 100% of 'NaN' outcomes.



Table 22: Results of 5-parameter estimation of a 3-type model I.

True $\beta, g_2, b_2, g_3, b_3$	(a) $\hat{\beta}$			(b) $\hat{g}_2$			(c) $\hat{b}_2$			(d) $\hat{g}_3$			(e) $\hat{b}_3$			
	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	NN
.5, .4, .3, .2, .15	.39	.45	.53	.37	.38	.27	.25	.28	.19	.24	.22	.27	.20	.17	.19	0%
3, .4, .3, .2, .15	2.53	2.77	1.66	.37	.38	.30	.25	.28	.19	.23	.21	.30	.20	.17	.19	0%
10, .4, .3, .2, .15	9.23	8.70	4.23	.34	.35	.25	.24	.26	.16	.25	.24	.25	.22	.19	.16	1%
.5, -.4, -.3, .2, .15	.46	.53	.50	-.53	-.53	.17	-.26	-.28	.11	.34	.33	.18	.11	.13	.11	0%
3, -.4, -.3, .2, .15	2.87	2.92	3.22	-.39	-.40	.16	-.24	-.27	.12	.23	.26	.19	.09	.12	.11	0%
10, -.4, -.3, .2, .15	11.16	11.10	10.62	-.35	-.35	.17	-.25	-.27	.12	.21	.28	.19	.10	.13	.11	0%
.5, -.4, .3, .2, -.15	.50	.55	.50	-.53	-.53	.17	.26	.28	.11	.34	.34	.17	-.11	-.13	.11	0%
3, -.4, .3, .2, -.15	2.76	2.84	3.32	-.38	-.40	.17	.24	.26	.11	.24	.27	.20	-.09	-.11	.11	0%
10, -.4, .3, .2, -.15	12.14	12.27	10.45	-.35	-.36	.17	.25	.28	.12	.20	.26	.18	-.10	-.14	.11	0%
.5, .8, .6, .4, .3	.51	.48	.19	.75	.78	.45	.45	.50	.33	.45	.42	.45	.45	.40	.33	25%
3, .8, .6, .4, .3	2.47	2.32	.90	.59	.65	.41	.44	.48	.29	.62	.55	.41	.45	.41	.29	45%
10, .8, .6, .4, .3	9.97	9.89	.75	.79	.63	.20	.59	.47	.15	.42	.58	.21	.31	.43	.15	89%
.5, -.8, -.6, .4, .3	.33	.43	.34	-1.00	-1.01	.30	-.69	-.71	.22	.61	.62	.31	.39	.41	.22	0%
3, -.8, -.6, .4, .3	2.29	3.15	2.49	-.85	-.91	.26	-.64	-.66	.22	.49	.54	.30	.34	.37	.21	9%
10, -.8, -.6, .4, .3	10.06	10.12	2.09	-.80	-.80	.06	-.60	-.60	.05	.40	.41	.09	.30	.30	.04	37%
.5, -.8, .6, .4, -.3	.34	.43	.34	-1.02	-1.01	.32	.69	.73	.23	.62	.61	.33	-.39	-.43	.23	0%
3, -.8, .6, .4, -.3	2.49	3.32	2.51	-.83	-.89	.26	.62	.66	.21	.45	.51	.28	-.33	-.36	.21	8%
10, -.8, .6, .4, -.3	10.03	10.05	2.67	-.80	-.79	.07	.60	.59	.06	.40	.41	.12	-.30	-.29	.05	37%
.5, 1.2, .9, .8, .6	.50	.43	.15	1.04	1.08	.60	.77	.82	.58	.96	.92	.60	.71	.68	.58	85%
3, 1.2, .9, .8, .6																100%
10, 1.2, .9, .8, .6																100%
.5, -1.2, -.9, .8, .6	.26	.32	.30	-1.58	-1.60	.46	-1.12	-1.14	.43	1.22	1.23	.49	.82	.84	.42	5%
3, -1.2, -.9, .8, .6	2.79	2.49	1.15	-1.24	-1.32	.24	-.91	-.88	.25	.84	1.02	.49	.61	.59	.23	62%
10, -1.2, -.9, .8, .6																100%
.5, -1.2, .9, .9, -.6	.24	.32	.31	-1.60	-1.60	.45	1.15	1.14	.43	1.22	1.23	.49	-.85	-.85	.43	5%
3, -1.2, .9, .9, -.6	2.74	2.38	1.18	-1.24	-1.34	.25	.91	.88	.26	.85	1.05	.51	-.61	-.59	.25	62%
10, -1.2, .9, .8, -.6																100%

Note: Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0, 0.1^2)$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD) are reported. Figures are rounded to 2 decimal digits. 'NN' column reports the percentage of runs with 'NaN' outcome rounded to integer numbers.

Table 23: Results of 5-parameter estimation of a 3-type model II.

True $\beta, g_2, b_2, g_3, b_3$	(a) $\hat{\beta}$			(b) $\hat{g}_2$			(c) $\hat{b}_2$			(d) $\hat{g}_3$			(e) $\hat{b}_3$			
	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	NN
.5, .4, .3, .2, .15	.39	.48	.35	.30	.33	.23	.25	.25	.19	.27	.25	.22	.22	.20	.18	0%
3, .4, .3, .2, .15	3.33	3.86	2.43	.35	.33	.16	.26	.26	.14	.23	.24	.16	.21	.18	.14	0%
10, .4, .3, .2, .15	10.61	12.69	8.15	.37	.31	.13	.27	.24	.12	.25	.25	.15	.20	.18	.12	3%
.5, -.4, -.3, .2, .15	.42	.57	.40	-.42	-.45	.14	-.32	-.34	.13	.21	.24	.15	.16	.18	.12	0%
3, -.4, -.3, .2, .15	3.56	3.53	3.17	-.35	-.31	.14	-.25	-.23	.12	.17	.15	.08	.12	.12	.07	0%
10, -.4, -.3, .2, .15	10.33	9.05	10.12	-.37	-.29	.17	-.27	-.22	.13	.18	.14	.09	.13	.12	.06	11%
.5, -.4, .3, .2, -.15	.41	.56	.40	-.43	-.45	.14	.32	.33	.13	.21	.24	.14	-.17	-.19	.12	0%
3, -.4, .3, .2, -.15	3.80	3.51	3.31	-.35	-.30	.14	.25	.22	.12	.17	.14	.08	-.12	-.12	.06	0%
10, -.4, .3, .2, -.15	10.28	9.37	10.02	-.37	-.30	.16	.27	.22	.13	.18	.15	.09	-.13	-.12	.06	12%
.5, .8, .6, .4, .3	.47	.45	.18	.72	.68	.34	.52	.52	.29	.49	.51	.34	.42	.38	.28	1%
3, .8, .6, .4, .3	2.94	2.99	.41	.46	.57	.21	.37	.42	.17	.77	.63	.21	.56	.47	.17	21%
10, .8, .6, .4, .3	9.41	9.68	2.83	.51	.56	.23	.37	.42	.20	.79	.64	.22	.58	.46	.19	70%
.5, -.8, -.6, .4, .3	.49	.52	.19	-.81	-.81	.09	-.60	-.61	.10	.40	.41	.08	.30	.31	.08	0%
3, -.8, -.6, .4, .3	2.94	2.43	1.76	-.79	-.70	.27	-.59	-.53	.19	.39	.35	.13	.30	.30	.05	7%
10, -.8, -.6, .4, .3	8.26	3.09	7.90	-.78	-.46	.40	-.56	-.35	.29	.38	.23	.20	.28	.27	.09	67%
.5, -.8, .6, .4, -.3	.49	.52	.18	-.81	-.81	.09	.60	.61	.10	.40	.41	.07	-.30	-.31	.08	0%
3, -.8, .6, .4, -.3	2.96	2.52	1.68	-.80	-.71	.25	.59	.54	.18	.39	.35	.13	-.29	-.29	.05	5%
10, -.8, .6, .4, -.3	8.56	4.07	7.65	-.79	-.50	.39	.58	.38	.28	.39	.25	.19	-.28	-.27	.09	69%
.5, 1.2, .9, .8, .6																100%
3, 1.2, .9, .8, .6																100%
10, 1.2, .9, .8, .6																100%
.5, -1.2, -.9, .8, .6	.51	.49	.14	-1.19	-1.16	.17	-.89	-.87	.14	.79	.77	.12	.59	.60	.07	86%
3, -1.2, -.9, .8, .6																100%
10, -1.2, -.9, .8, .6																100%
.5, -1.2, .9, .8, -.6	.51	.49	.12	-1.19	-1.16	.18	.89	.87	.14	.79	.77	.12	-.60	-.62	.13	87%
3, -1.2, .9, .8, -.6																100%
10, -1.2, .9, .8, -.6																100%

Note: Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from normal distribution  $N(0, 1)$ ,  $R = 1.0001$ . Each sample is based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Sample medians, means, standard deviations (SD) are reported. Figures are rounded to 2 decimal digits. 'NN' column reports the percentage of runs with 'NaN' outcome rounded to integer numbers.

Table 24: Empirical results of the 2-type *fraction* model estimation

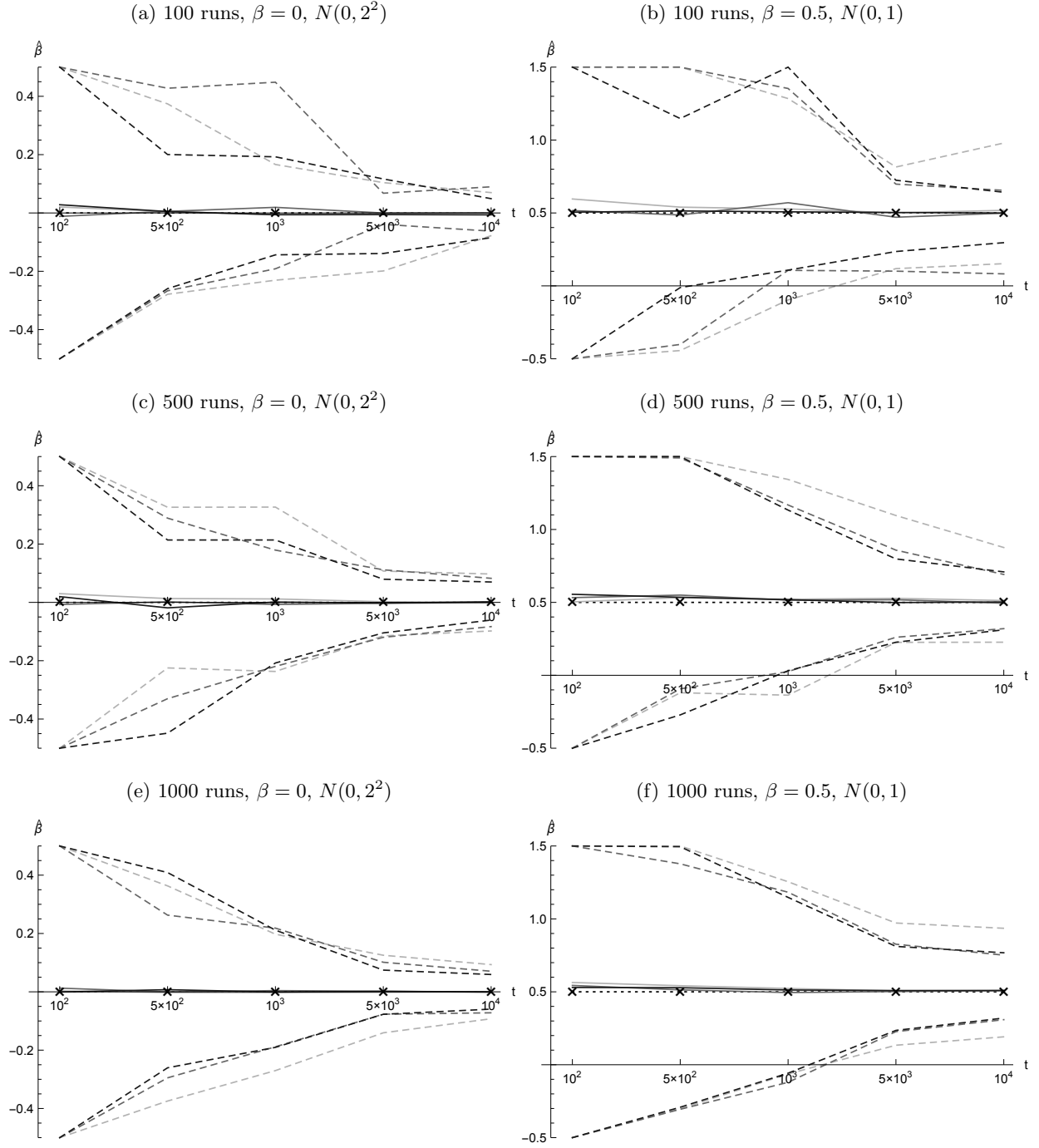
Data, MA period	(a) $\widehat{fraction}$			(b) $\widehat{g_2}$			(c) $\widehat{b_2}$			(d) $\widehat{noise\ intensity}$			(e) $LL$		
	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD	Med.	Mean	SD
SP500, 61	.569	.555	.088	1.963	1.986	.362	.006	.003	.121	.559	.558	.044	-4.022	-4.023	.013
NASDAQ, 61	.555	.541	.086	1.972	1.985	.344	-.001	-.001	.117	.565	.566	.054	-5.118	-5.119	.019
DAX, 61	.559	.547	.090	2.006	2.021	.387	.008	.005	.123	.489	.490	.035	-5.733	-5.733	.011
FTSE, 61	.556	.545	.085	1.980	1.997	.357	-.009	-.007	.121	.519	.520	.037	-5.507	-5.507	.011
HSI, 61	.559	.549	.082	1.993	2.020	.344	-.006	-.001	.120	.531	.531	.045	-6.999	-7.000	.016
N225, 61	.562	.553	.084	1.995	2.023	.361	.005	.001	.124	.482	.484	.032	-6.700	-6.701	.010
SP500, 241	.562	.556	.084	2.161	2.228	.425	.002	-.001	.133	.331	.348	.053	-4.073	-4.088	.049
NASDAQ, 241	.562	.556	.086	2.191	2.244	.434	-.002	-.001	.133	.311	.345	.077	-5.195	-5.231	.089
DAX, 241	.565	.559	.088	2.219	2.275	.455	.013	.003	.137	.256	.279	.069	-5.782	-5.814	.086
FTSE, 241	.562	.554	.087	2.179	2.227	.445	-.002	.003	.133	.322	.339	.053	-5.547	-5.563	.045
HSI, 241	.562	.554	.084	2.204	2.253	.446	-.009	-.006	.137	.276	.304	.070	-7.057	-7.095	.091
N225, 241	.570	.562	.087	2.232	2.281	.460	-.012	-.005	.138	.255	.272	.058	-6.744	-6.767	.069
Robustness check															
SP500 monthly, 13	.606	.592	.114	1.429	1.451	.306	-.000	.000	.116	.806	.801	.064	-5.150	-5.153	.014
SP500 weekly, 13	.635	.616	.128	1.404	1.422	.309	-.011	-.004	.115	.853	.848	.063	-4.598	-4.603	.028
SP500 weekly, 49	.485	.483	.093	1.591	1.624	.292	.012	.006	.117	.606	.606	.057	-4.721	-4.722	.017
SP500 R=1.001, 61	.566	.553	.089	1.967	1.984	.368	-.004	-.004	.121	.557	.557	.044	-4.022	-4.022	.013
SP500 R=1.001, 241	.570	.557	.085	2.220	2.239	.430	.003	-.001	.132	.332	.349	.054	-4.074	-4.089	.047

*Note:* Results are based on 1000 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 500$ . Sample medians, means, and standard deviations (SD) are reported. ‘LL’ denotes log-likelihoods of estimated models representing statistical fits. Figures are rounded to 3 decimal digits.

## **Appendix 2: Supplementary figures**

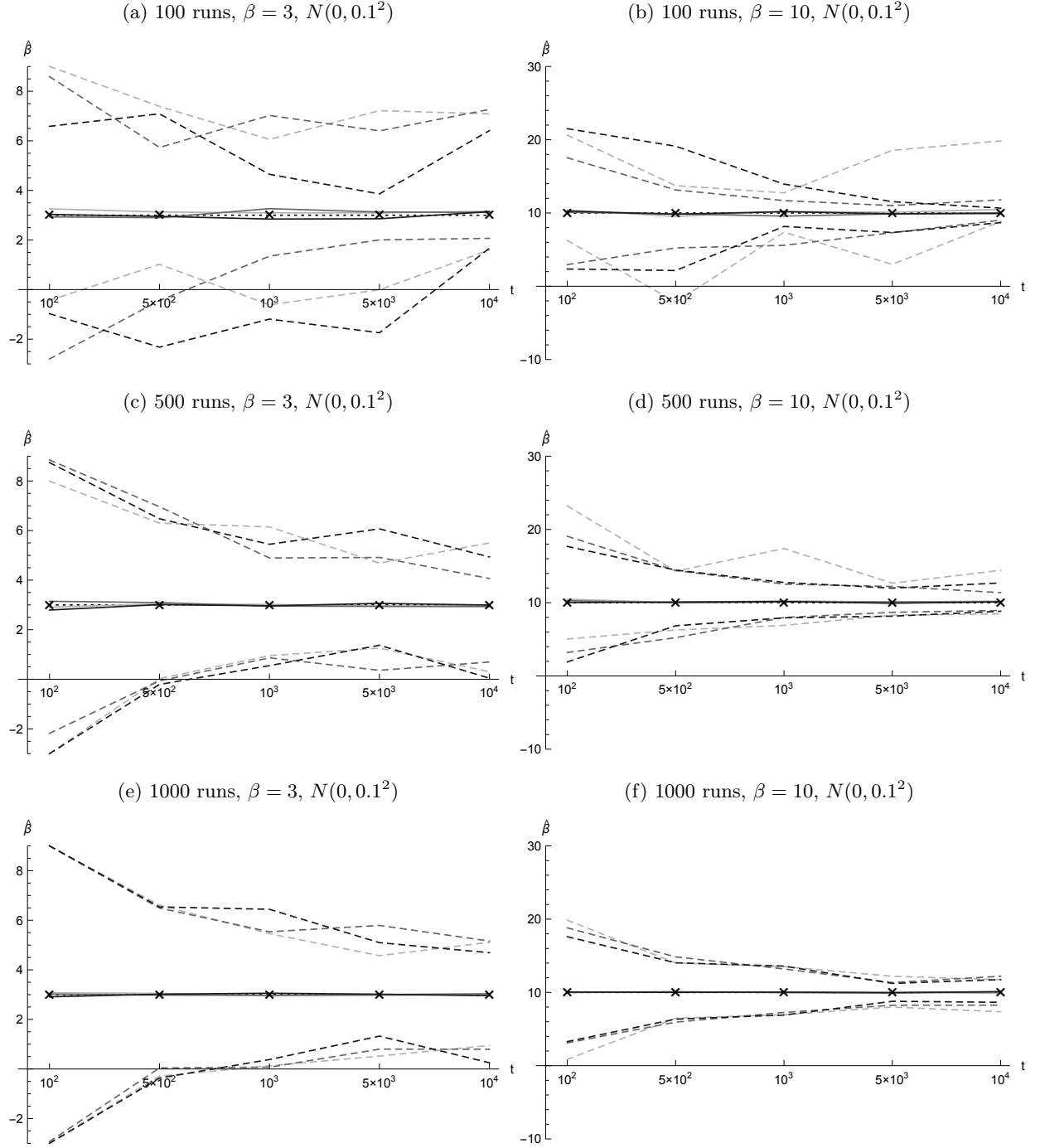
On the following pages, supplementary figures are provided.

Figure 7: Simulation results for various number of runs and  $\beta$ s I.



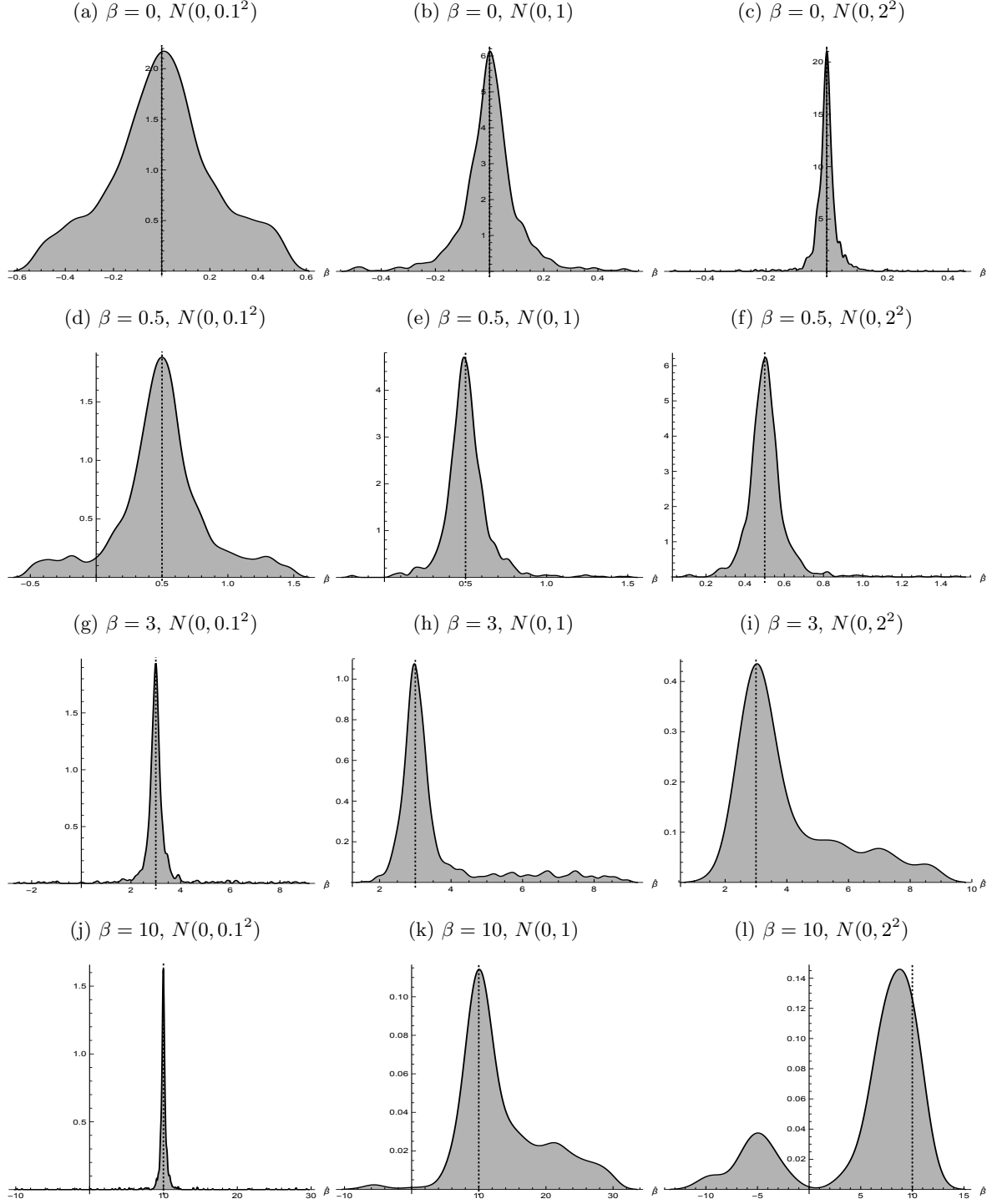
*Note:* Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distributions. Black dotted lines with  $\times$  depict the true  $\beta$ . Grey full lines depict sample means of estimated  $\hat{\beta}$ . Grey dashed lines depict 2.5% and 97.5% quantiles. Light grey colour represents results for  $N = 100$ , normal grey for  $N = 500$ , and dark grey for  $N = 1000$ . 't' (horizontal axis) stands for the length of generated time series.

Figure 8: Simulation results for various number of runs and  $\beta$ s II.



*Note:* Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distributions. Black dotted lines with  $\times$  depict the true  $\beta$ . Grey full lines depict sample means of estimated  $\beta$ . Grey dashed lines depict 2.5% and 97.5% quantiles. Light grey colour represents results for  $N = 100$ , normal grey for  $N = 500$ , and dark grey for  $N = 1000$ . 't' (horizontal axis) stands for the length of generated time series.

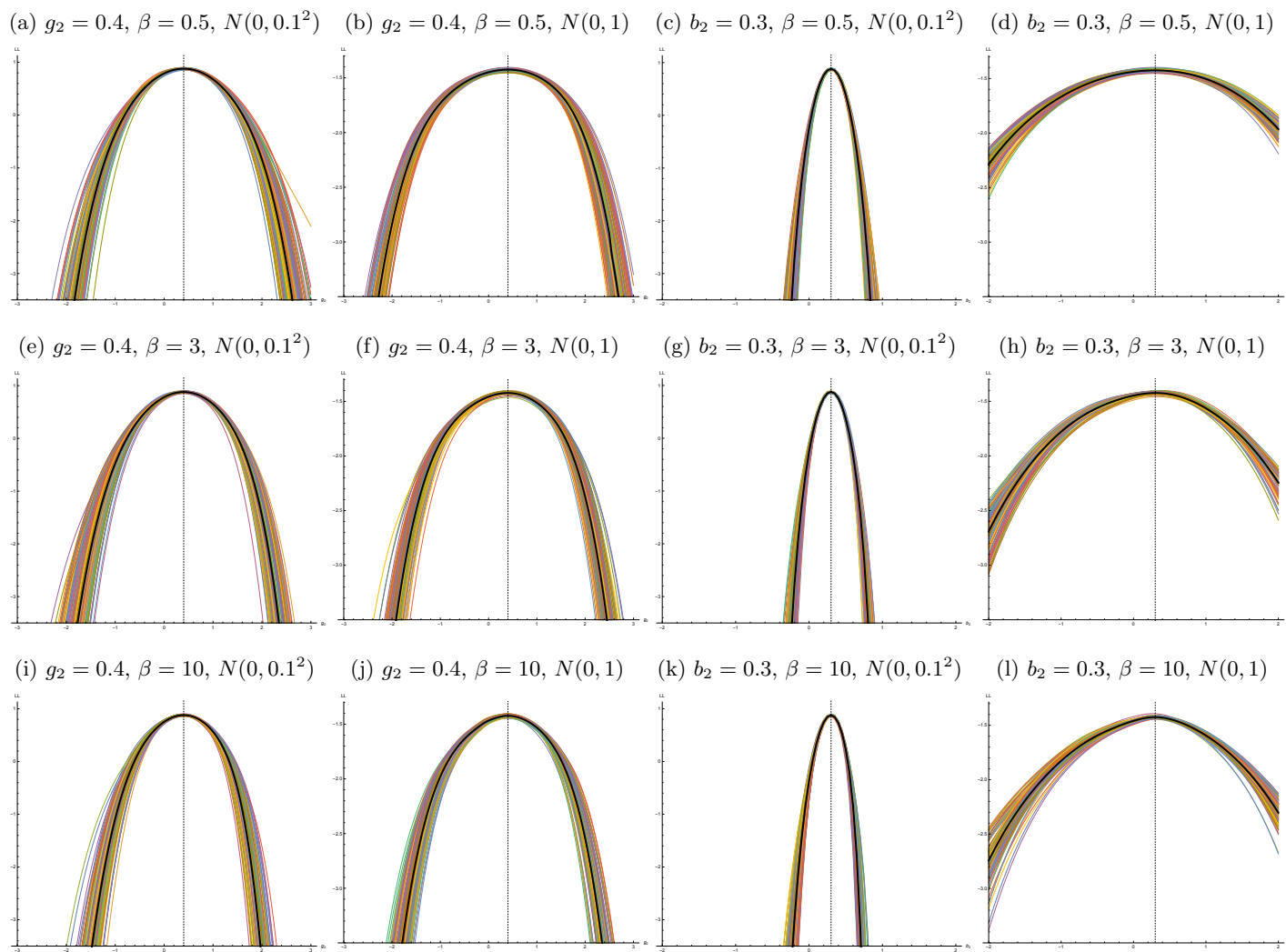
Figure 9: Smooth histograms for selected  $\hat{\beta}s$



*Note:* Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distributions,  $R = 1.0001$ . Each sample is based on 1000 random runs,  $H = 5$  possible trading strategies, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Black dotted lines depict the true  $\beta$ s. Produced using automatic `SmoothHistogram` kernel approximation function in *Wolfram Mathematica*.

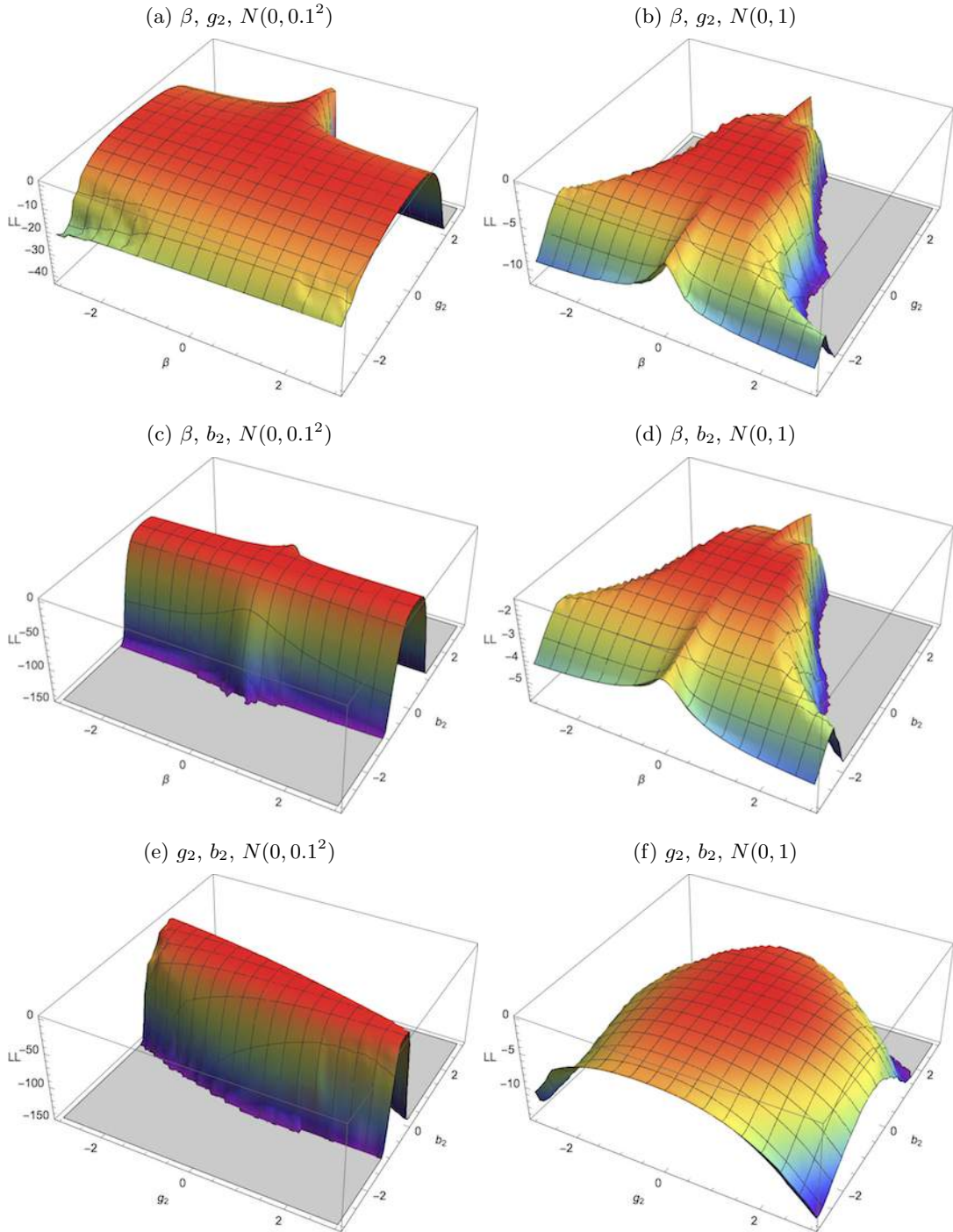


Figure 10: Simulated sub-log-likelihood fcns. for  $g_2$  and  $b_2$  estimation



*Note:* Results based on 100 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . Stochastic noise  $\epsilon_t$  and  $\{\varepsilon_i\}_{i=1}^N$  drawn from given normal distribution. Black dotted vertical lines depict the true  $g_2 = 0.4$  and  $b_2 = 0.3$ . Bold black full lines depict sample averages.

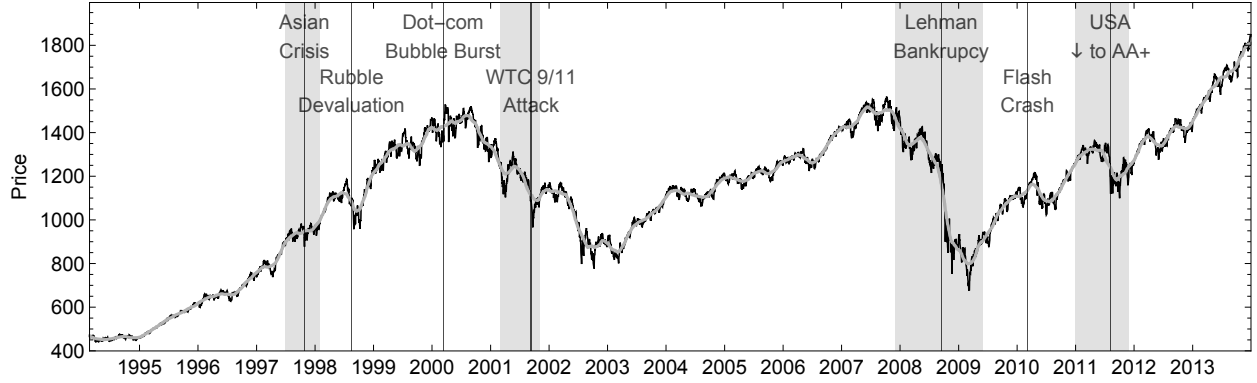
Figure 11: Simulated sub-log-likelihood functions in 3D



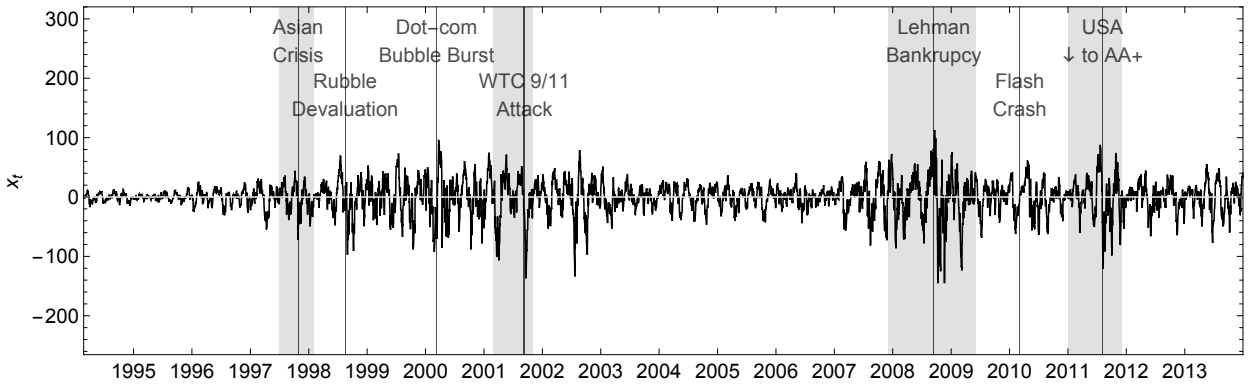
*Note:* Results averaged over 30 random runs, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ . The complete set of true parameters:  $\beta = 0.5$ ,  $g_2 = 0.4$ ,  $b_2 = 0.3$ . Stochastic noise  $\epsilon_t$  and  $\{\epsilon_i\}_{i=1}^N$  drawn from given normal distribution.

Figure 12: S&P500 fundamental price MA61 approximation

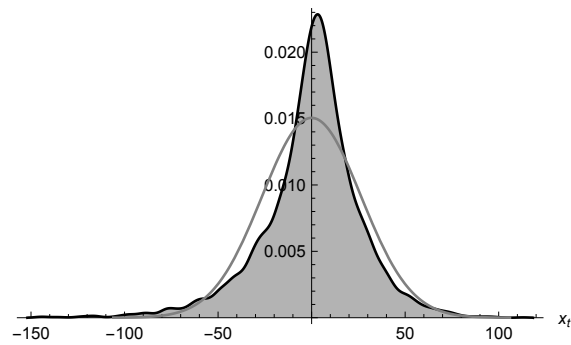
(a) MA window 61 days



(b)  $x_t$  time series



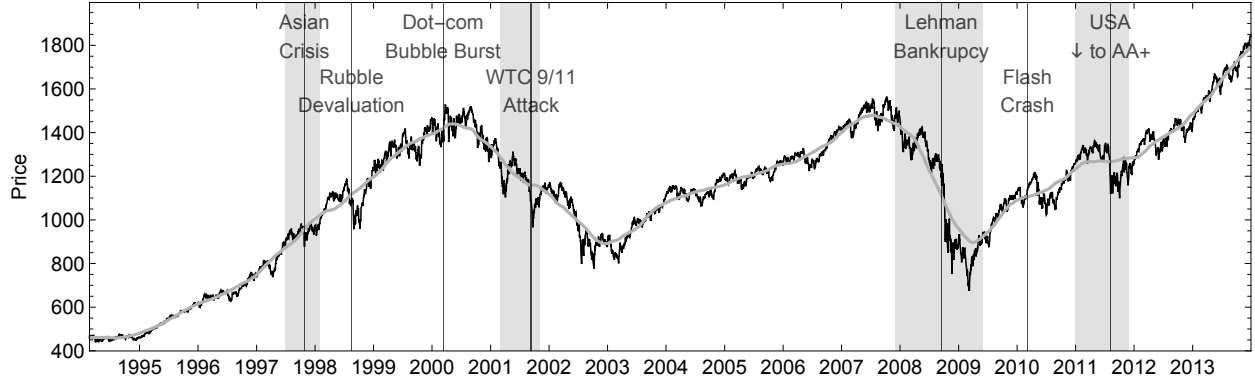
(c)  $x_t$  histogram



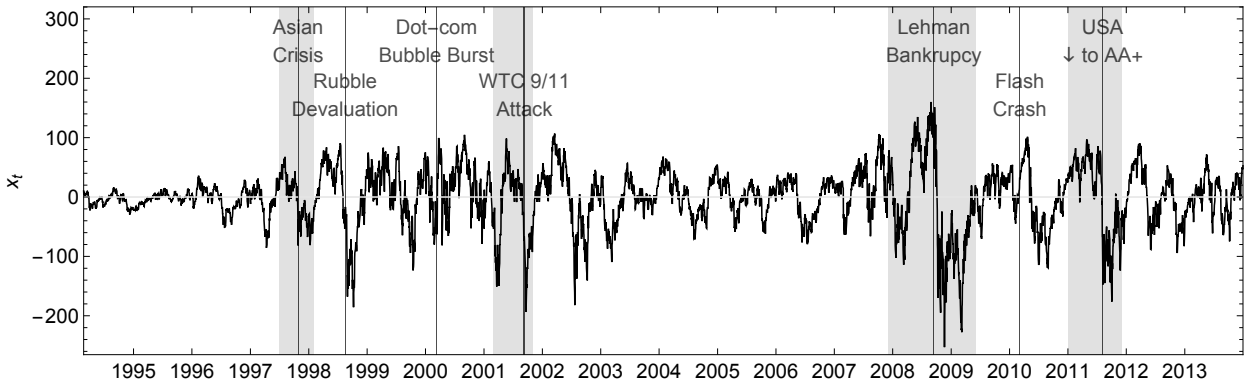
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 61 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 13: S&P500 fundamental price MA241 approximation

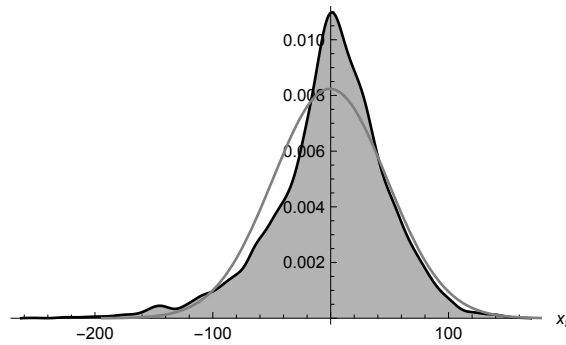
(a) MA window 241 days



(b)  $x_t$  time series



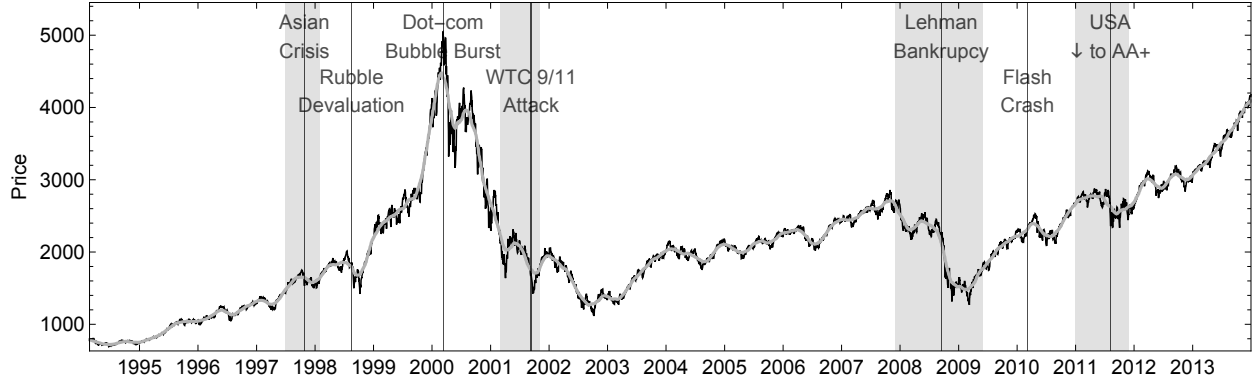
(c)  $x_t$  histogram



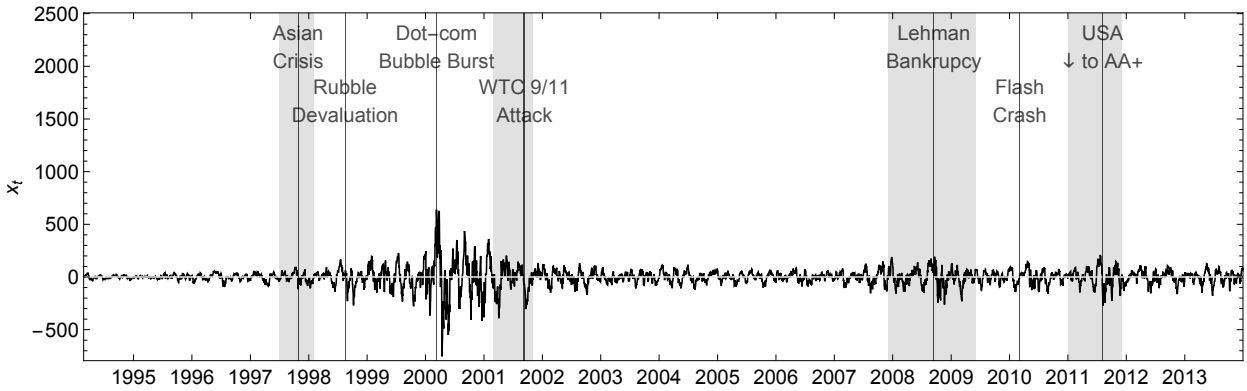
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 241 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 14: NASDAQ fundamental price MA61 approximation

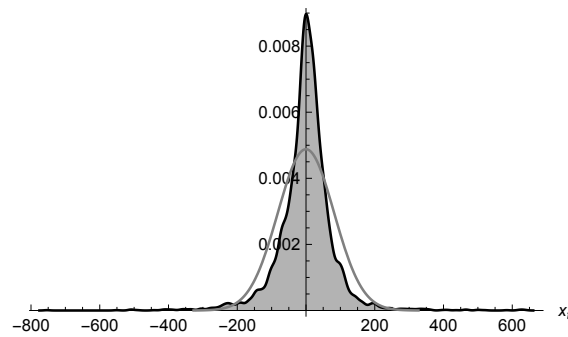
(a) MA window 61 days



(b)  $x_t$  time series



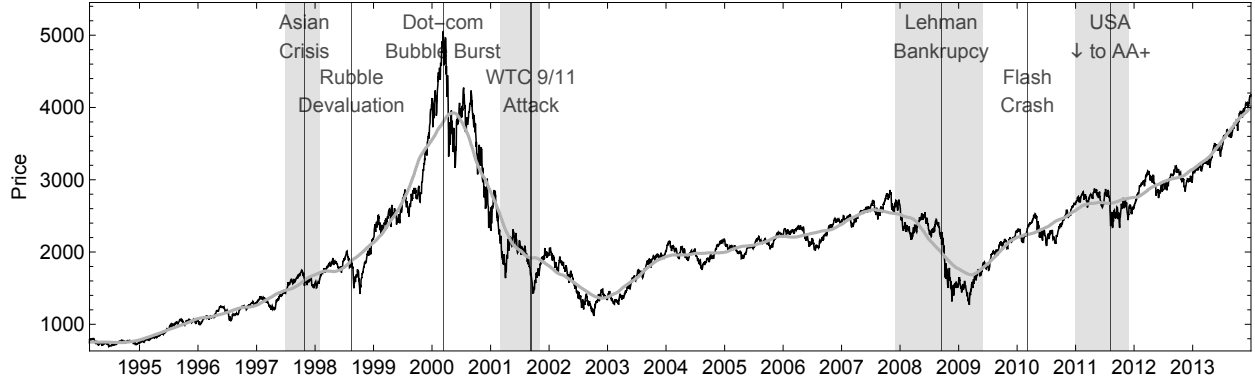
(c)  $x_t$  histogram



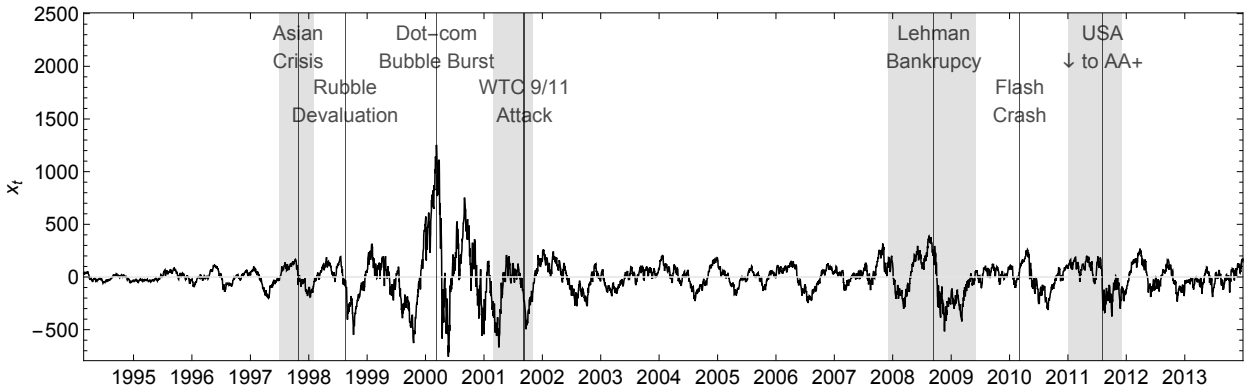
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 61 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 15: NASDAQ fundamental price MA241 approximation

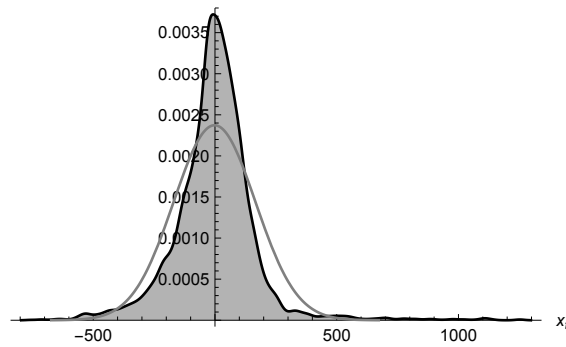
(a) MA window 241 days



(b)  $x_t$  time series



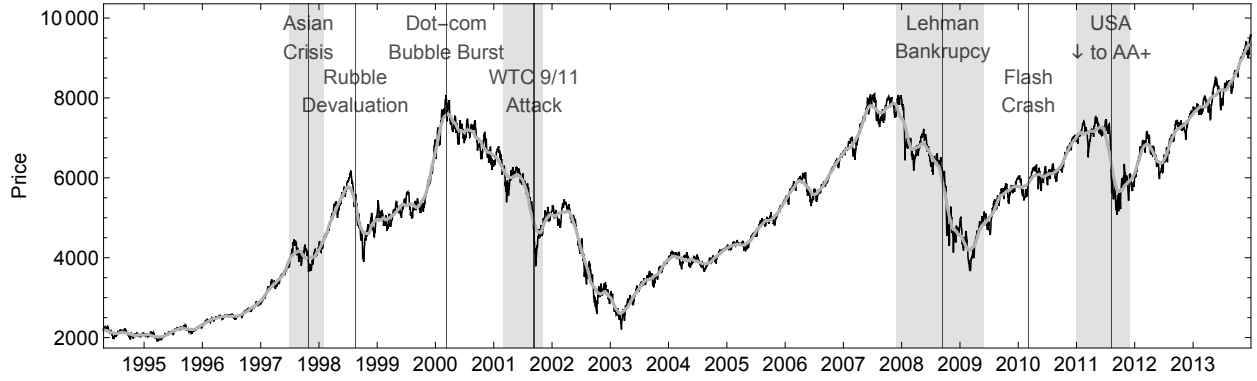
(c)  $x_t$  histogram



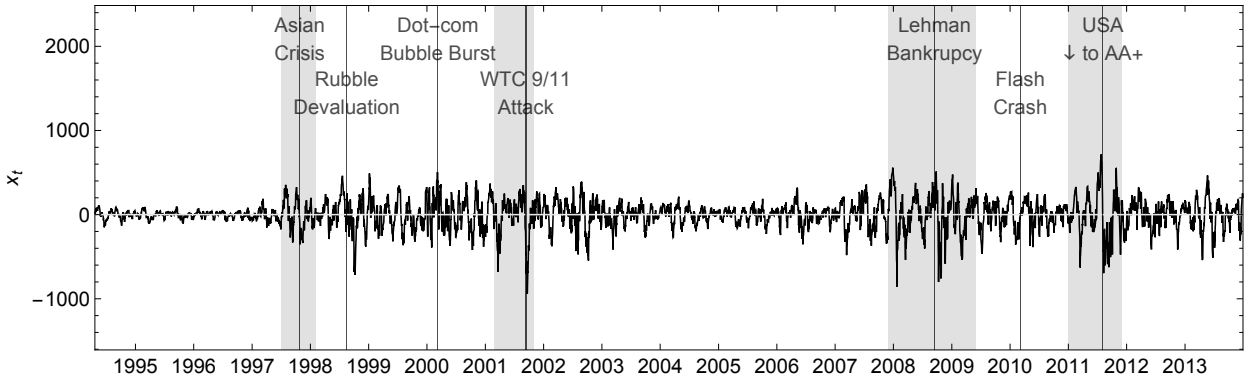
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 241 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 16: DAX fundamental price MA61 approximation

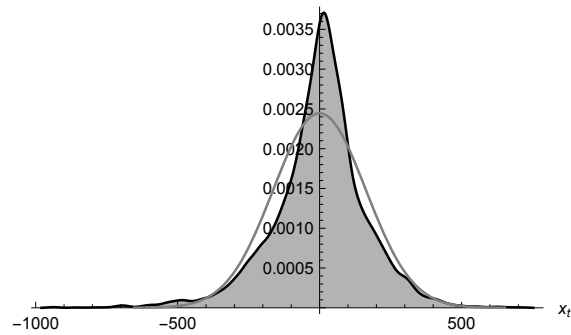
(a) MA window 61 days



(b)  $x_t$  time series



(c)  $x_t$  histogram

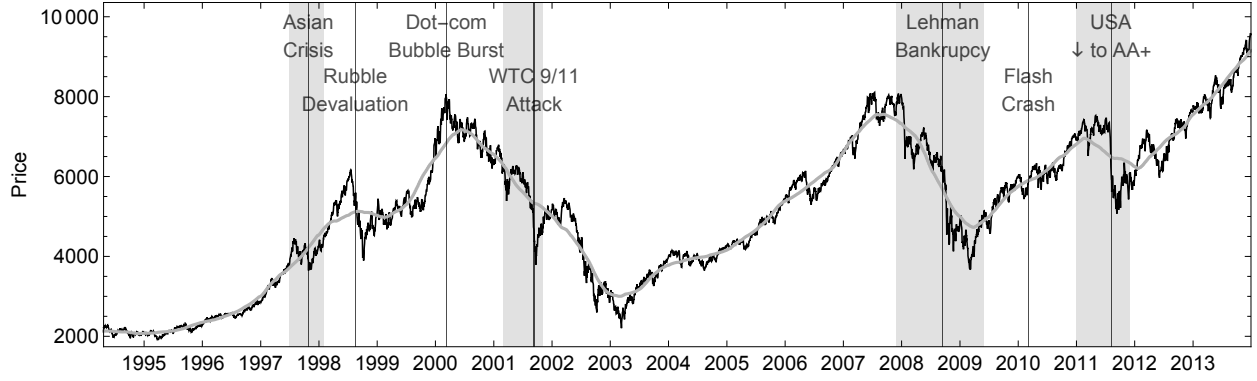


*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 61 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

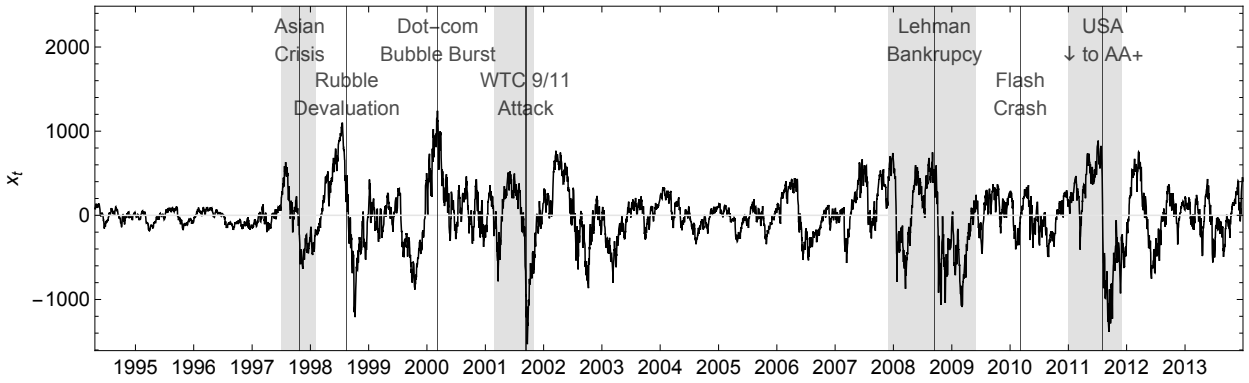


Figure 17: DAX fundamental price MA241 approximation

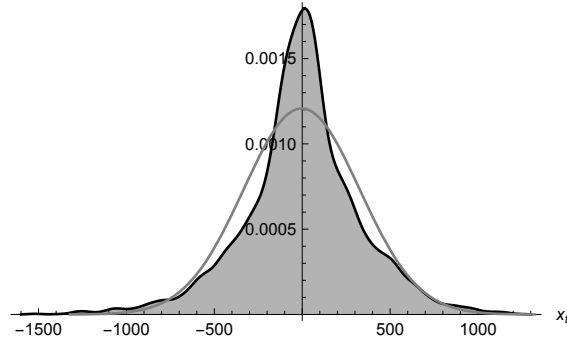
(a) MA window 241 days



(b)  $x_t$  time series



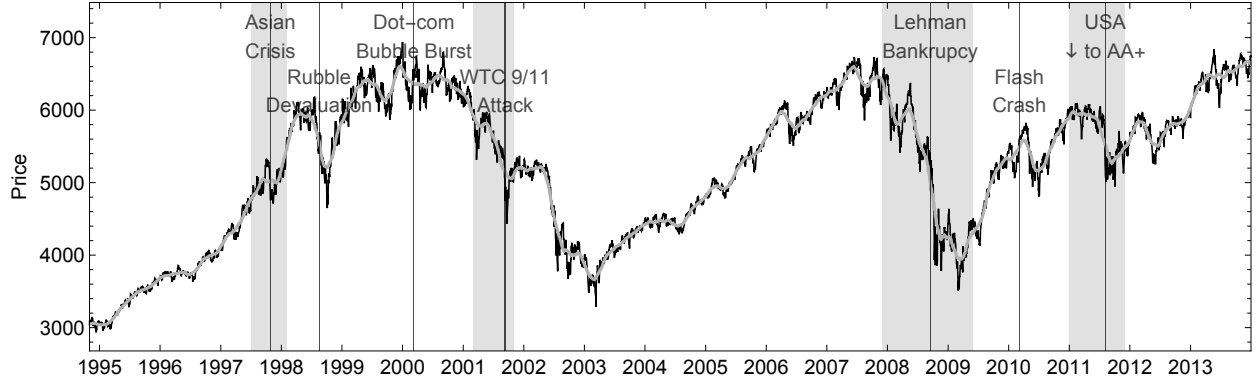
(c)  $x_t$  histogram



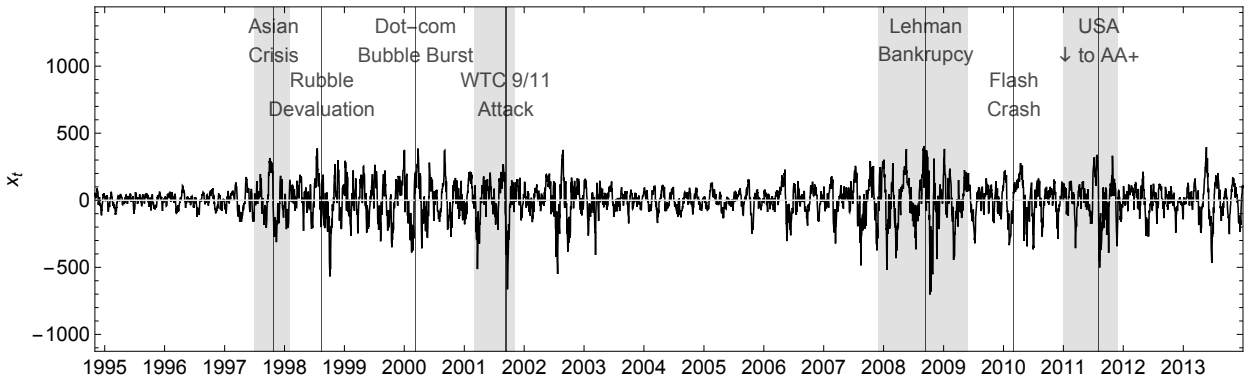
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 241 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 18: FTSE fundamental price MA61 approximation

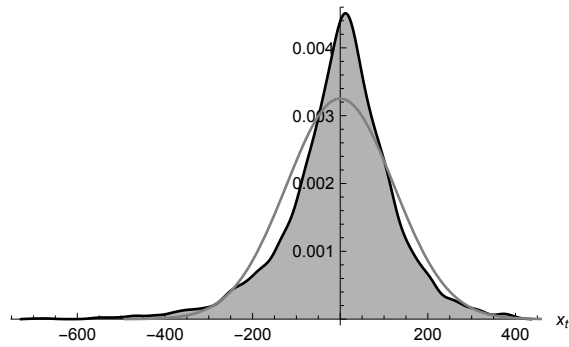
(a) MA window 61 days



(b)  $x_t$  time series



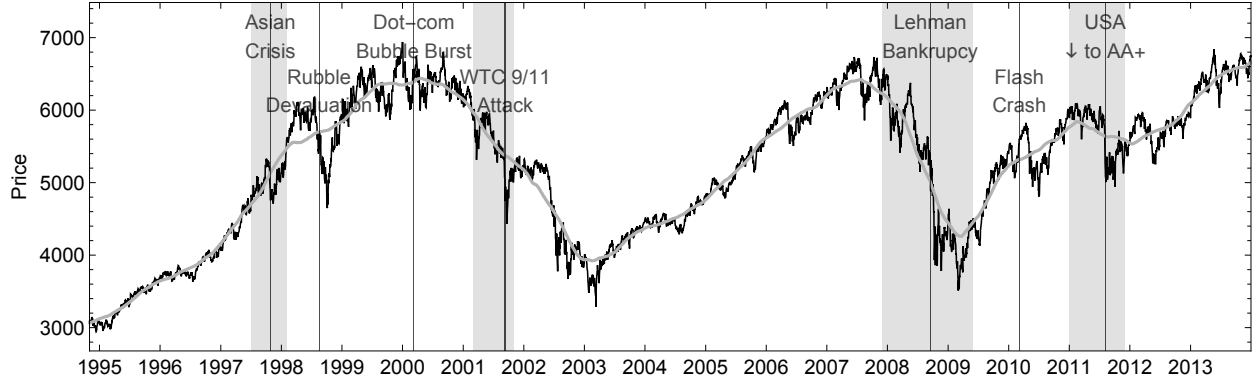
(c)  $x_t$  histogram



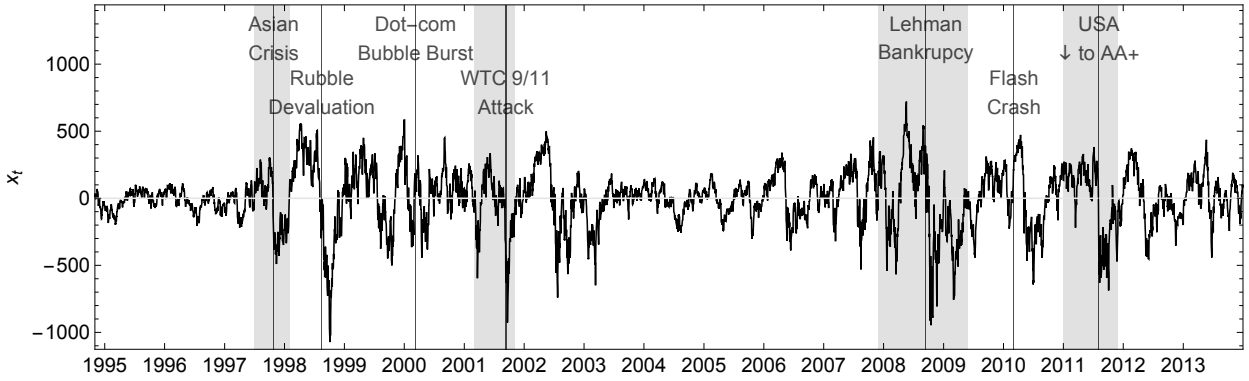
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 61 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 19: FTSE fundamental price MA241 approximation

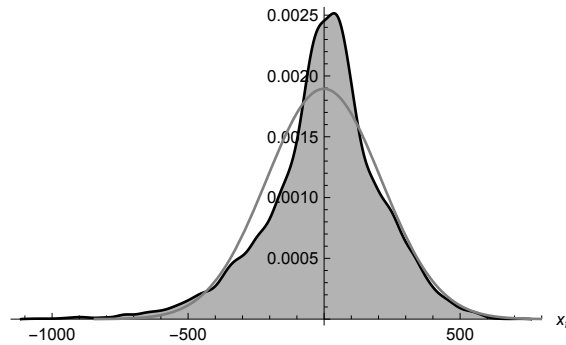
(a) MA window 241 days



(b)  $x_t$  time series



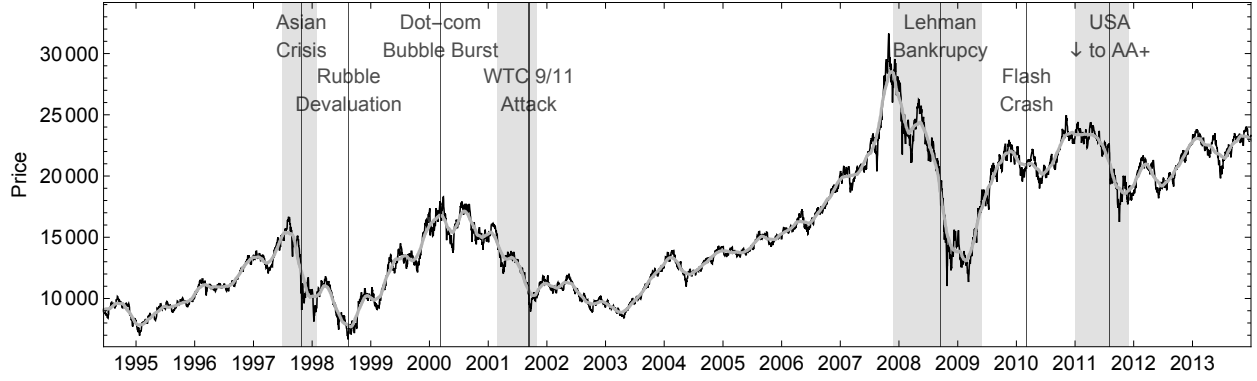
(c)  $x_t$  histogram



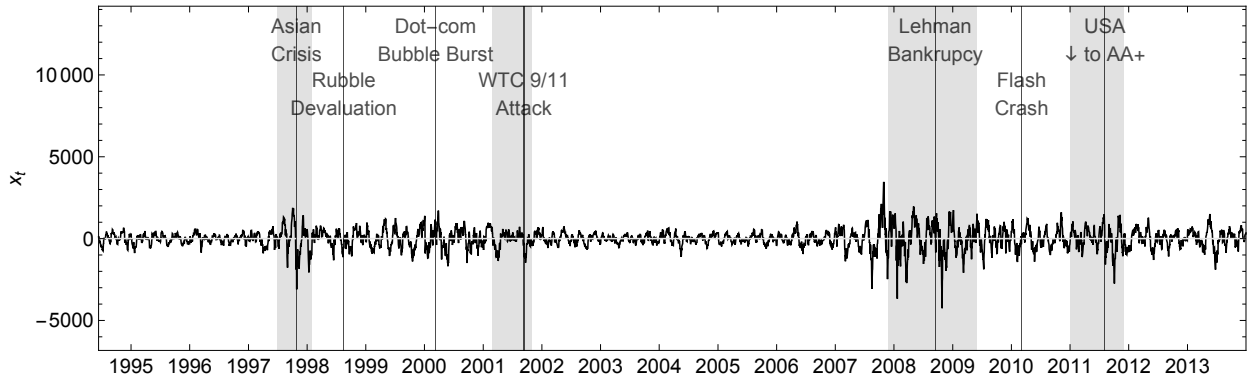
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 241 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 20: HSI fundamental price MA61 approximation

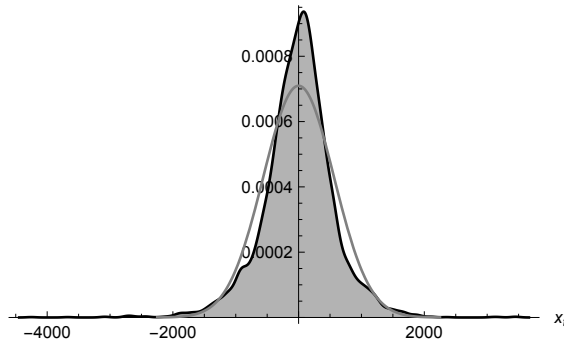
(a) MA window 61 days



(b)  $x_t$  time series



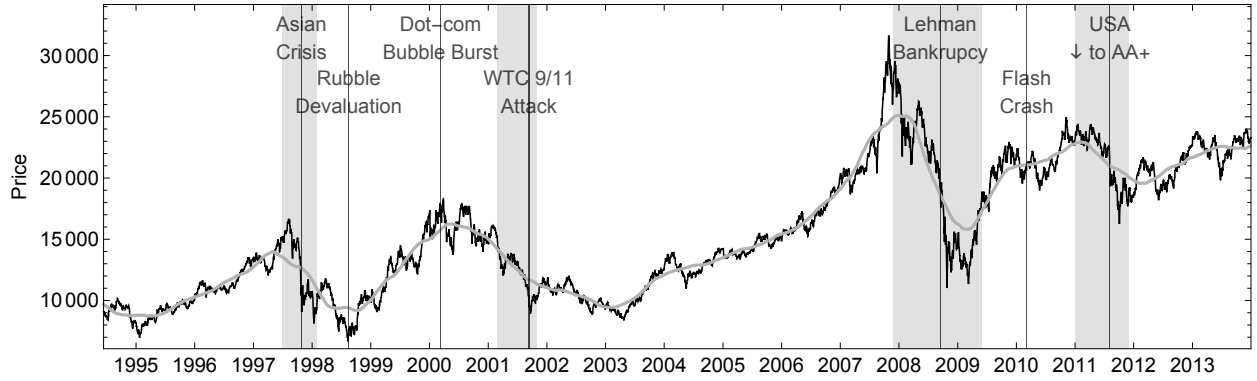
(c)  $x_t$  histogram



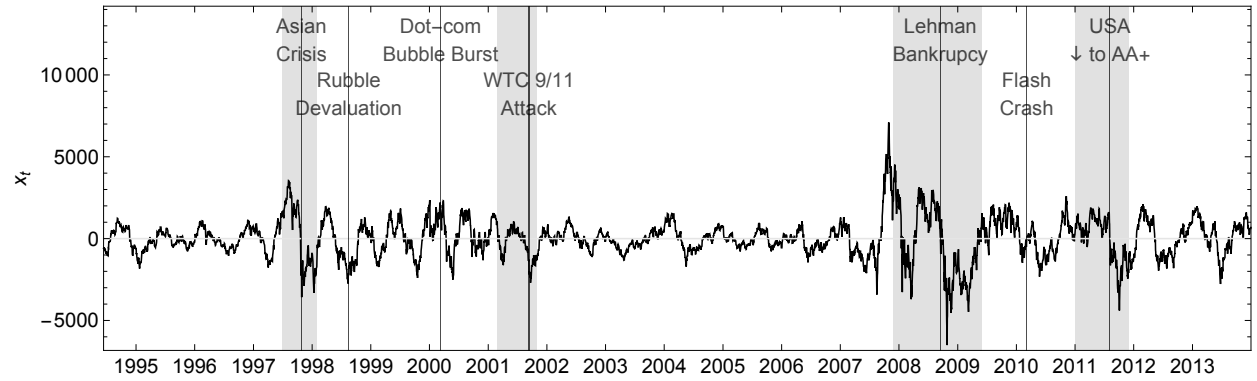
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 61 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 21: HSI fundamental price MA241 approximation

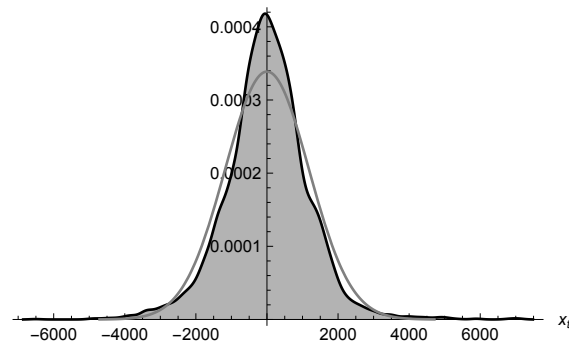
(a) MA window 241 days



(b)  $x_t$  time series



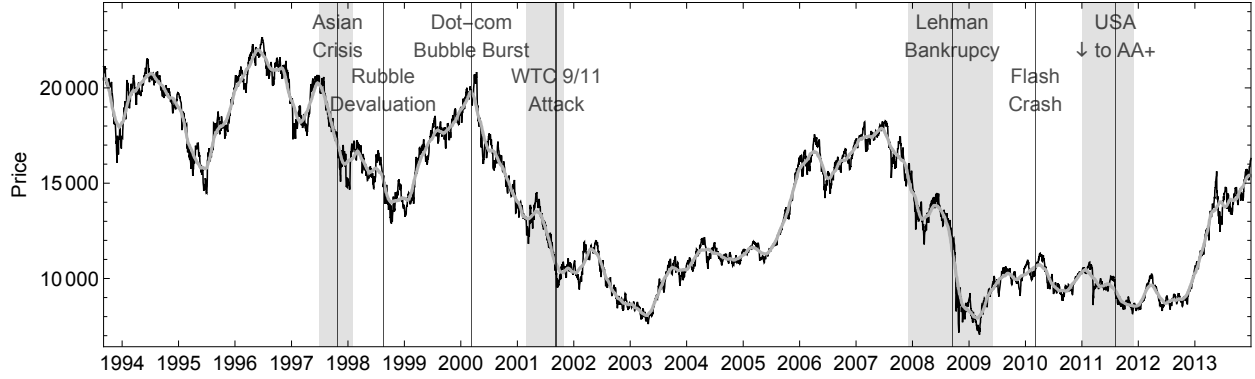
(c)  $x_t$  histogram



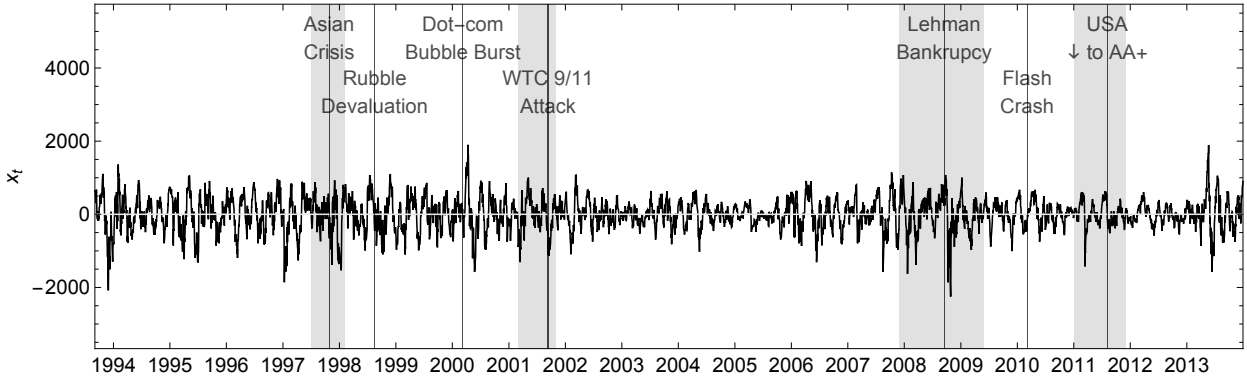
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 241 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 22: NIKKEI 225 fundamental price MA61 approximation

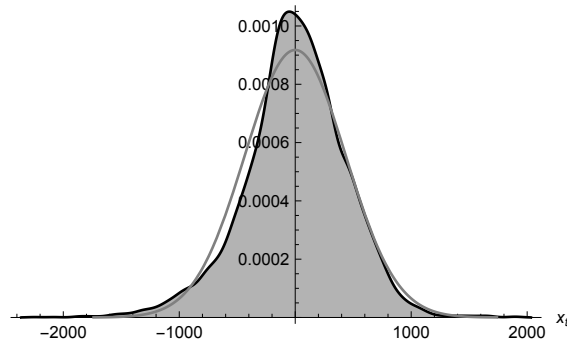
(a) MA window 61 days



(b)  $x_t$  time series



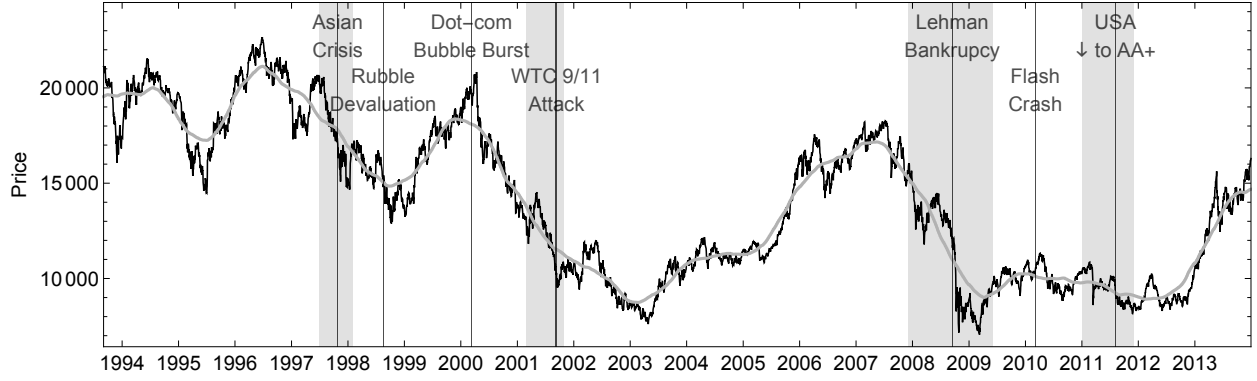
(c)  $x_t$  histogram



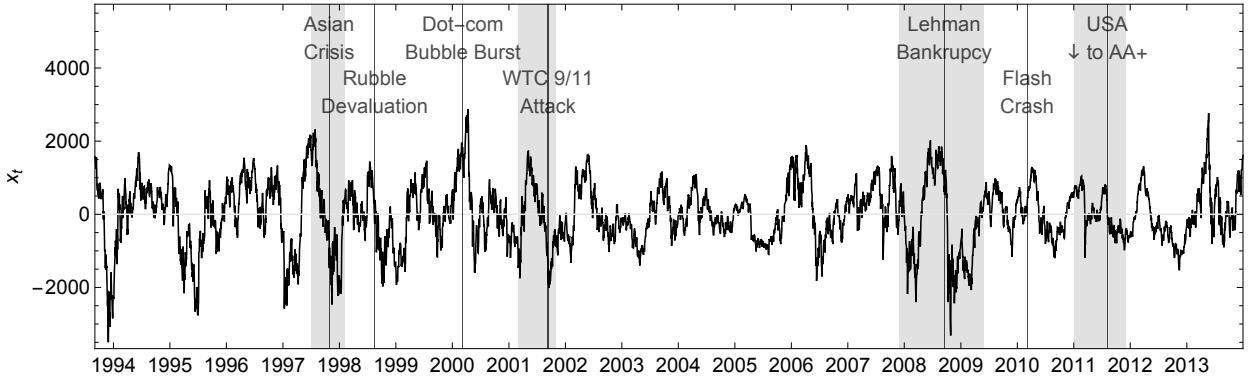
*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 61 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

Figure 23: NIKKEI 225 fundamental price MA241 approximation

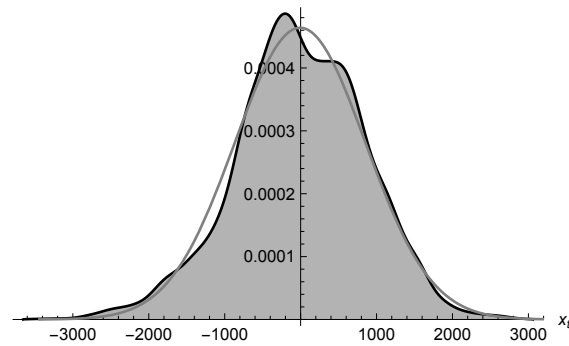
(a) MA window 241 days



(b)  $x_t$  time series



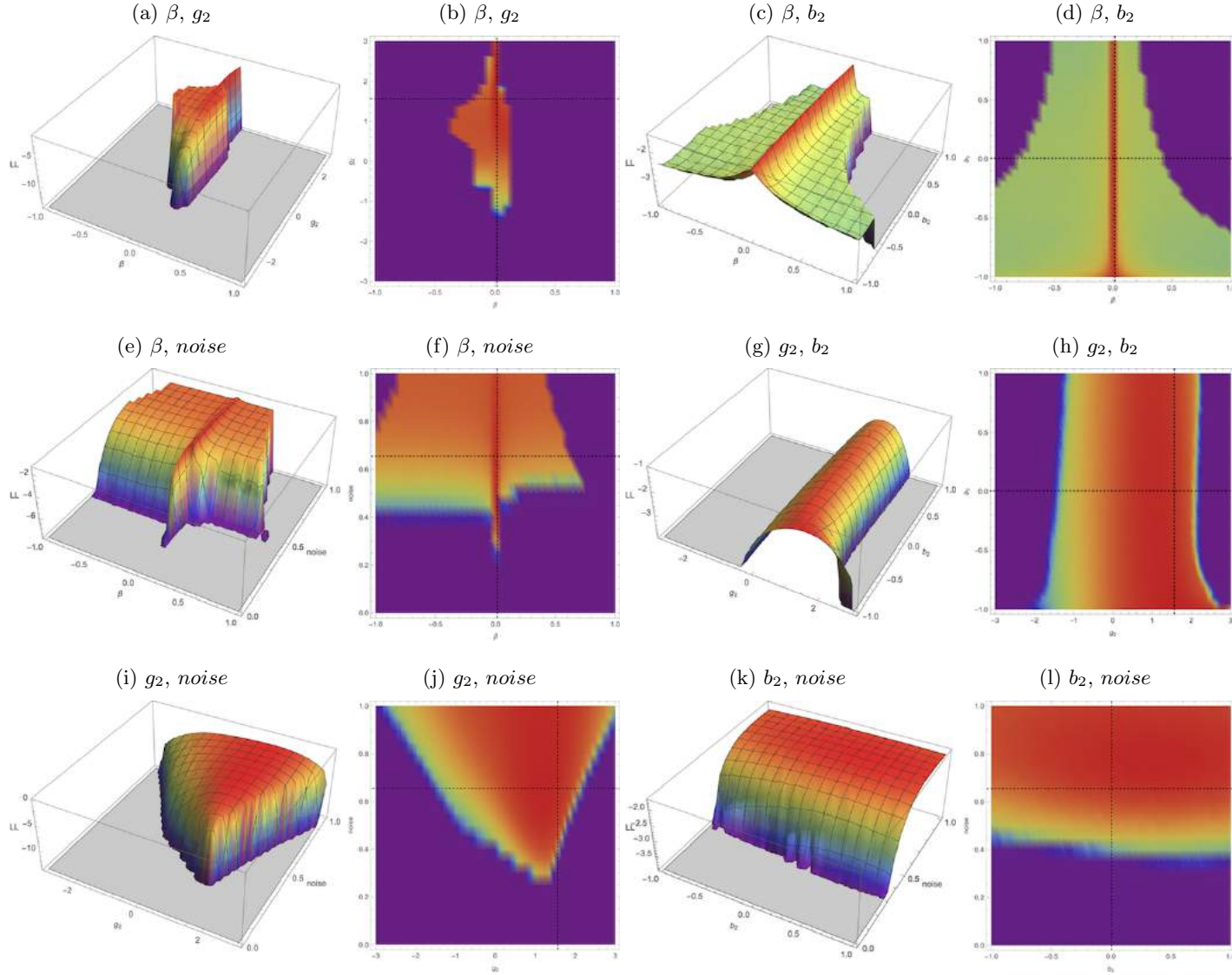
(c)  $x_t$  histogram



*Note:* (a) depicts the original price  $p_t$  in blacks and the fundamental price  $p_t^*$  approximation via 241 days centred MA in light grey. (b) plots the implied  $x_t = p_t - p_t^*$ . The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey). (c) shows the same data as (b) in a smooth histogram kernel approximation format in black together with the fit of  $N(\mu, \sigma^2)$  in grey.

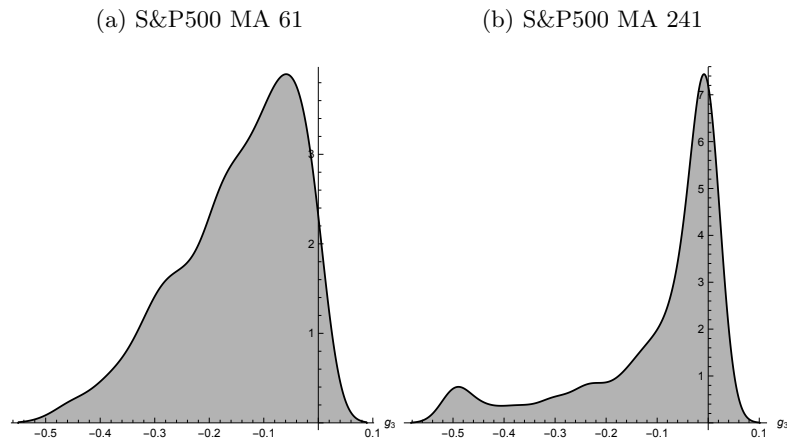


Figure 24: Simulated sub-log-likelihood functions in 3D



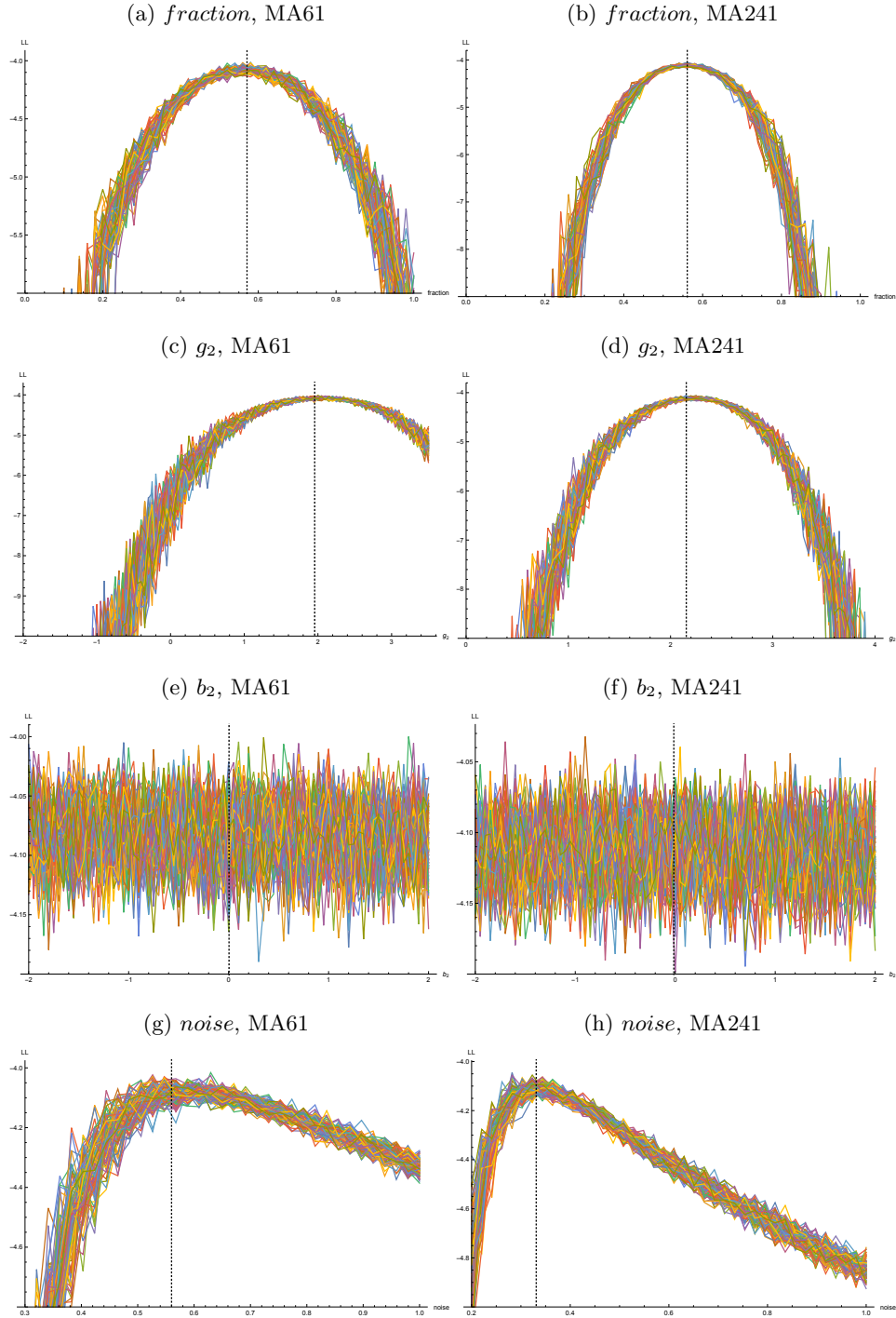
*Note:* Results averaged over 100 random runs, S&P500 data, MA61 fundamental price approximation, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ .  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distribution. Black dotted lines in horizontal projections depict the true values of parameters.

Figure 25: Smooth histogram of the contrarian coefficient  $g_3$



*Note:* Results are based on 500 random runs. Produced using automatic `SmoothHistogram` kernel approximation function in *Wolfram Mathematica*.

Figure 26: Simulated sub-log-likelihood fcn. for single parameters

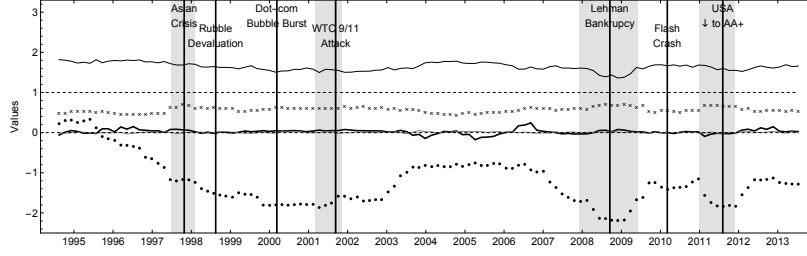


*Note:* Results based on 100 random runs, S&P500 data, given MA fundamental price approximation, number of observations  $t = 5000$ , and the kernel estimation precision  $N = 1000$ .  $\{\varepsilon_i\}_{i=1}^N$  drawn from normal distribution. Black dotted vertical lines depict estimated parameters (see Table 24).

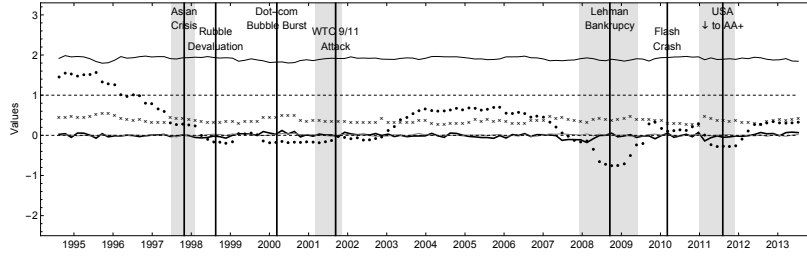
### Appendix 3: Rolling HAM estimates

Figure 27: Rolling estimates of the 2-type model for S&P500

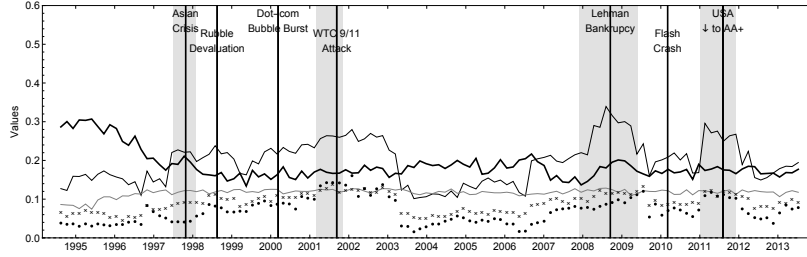
(a) MA61 fundamental price approximation



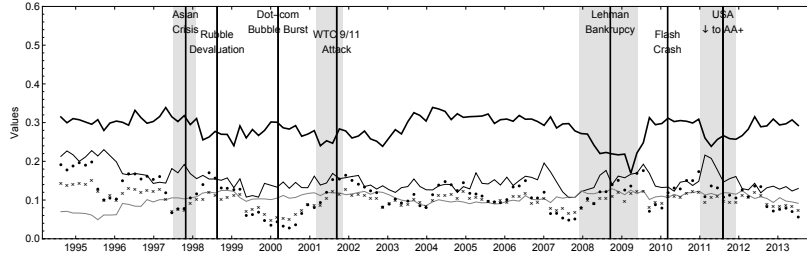
(b) MA241 fundamental price approximation



(c) MA61—related standard deviations



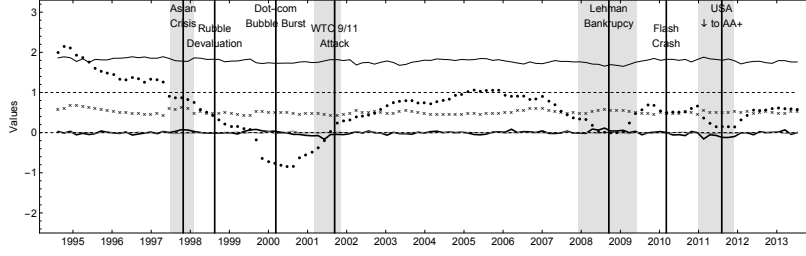
(d) MA241—related standard deviations



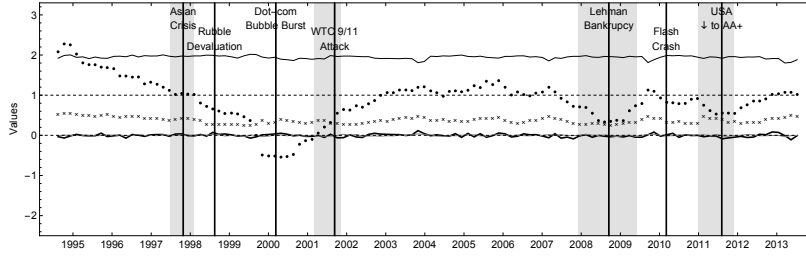
*Note:* Bold black full line depicts  $\hat{\beta}$ , black full line depicts  $\hat{g}_2$ , and grey full line depicts  $\hat{b}_2$ . *noise intensity* and *LL* are represented by  $\times$  and  $\bullet$ , respectively. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

Figure 28: Rolling estimates of the 2-type  $\beta$  model for NASDAQ

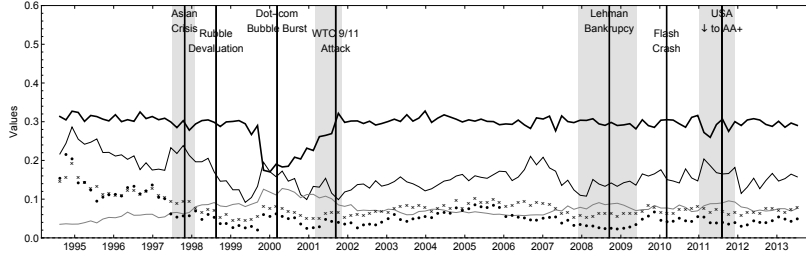
(a) MA61 fundamental price approximation



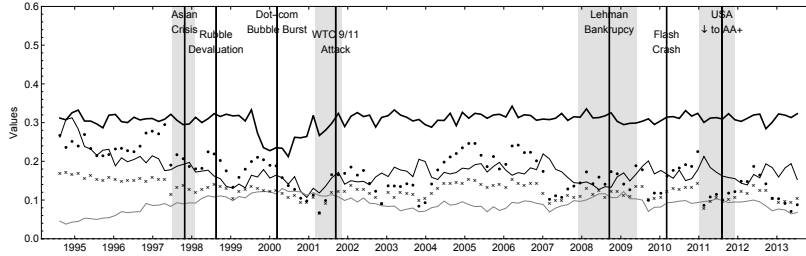
(b) MA241 fundamental price approximation



(c) MA61—related standard deviations



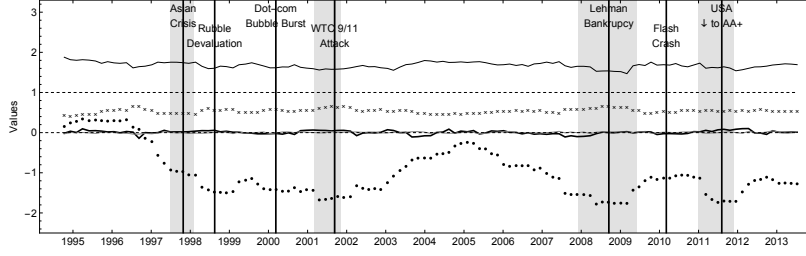
(d) MA241—related standard deviations



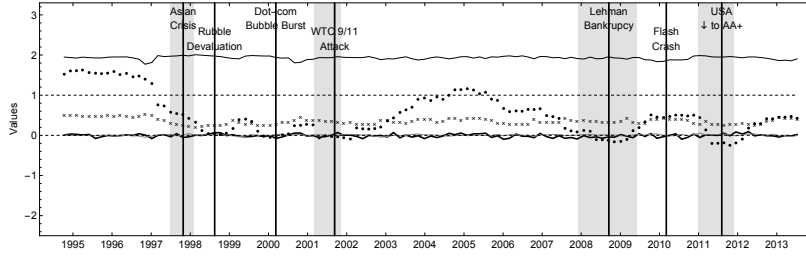
*Note:* Bold black full line depicts  $\hat{\beta}$ , black full line depicts  $\hat{g}_2$ , and grey full line depicts  $\hat{b}_2$ . *noise intensity* and *LL* are represented by  $\times$  and  $\bullet$ , respectively. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

Figure 29: Rolling estimates of the 2-type  $\beta$  model for DAX

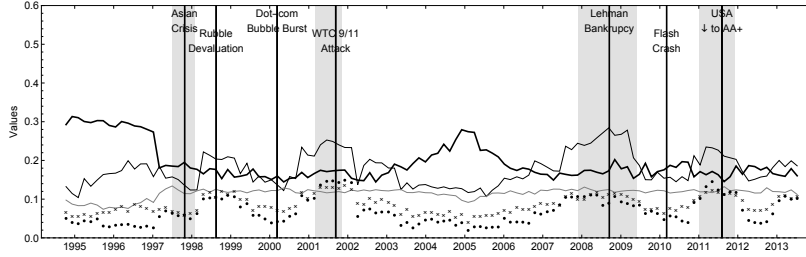
(a) MA61 fundamental price approximation



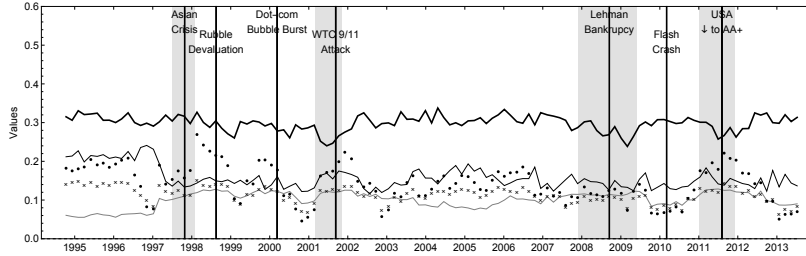
(b) MA241 fundamental price approximation



(c) MA61—related standard deviations



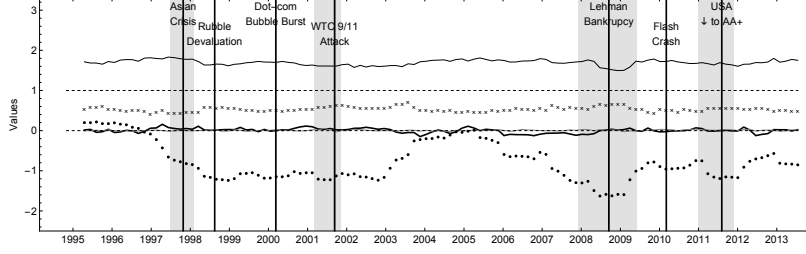
(d) MA241—related standard deviations



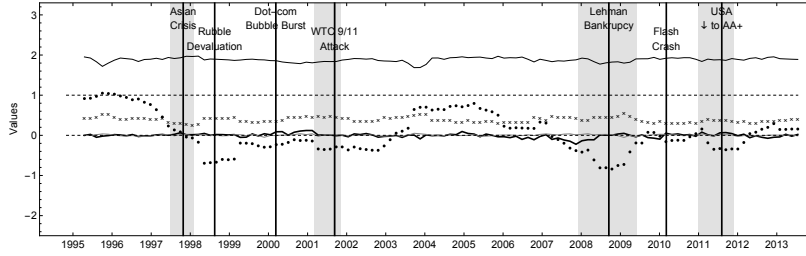
*Note:* Bold black full line depicts  $\hat{\beta}$ , black full line depicts  $\hat{g}_2$ , and grey full line depicts  $\hat{b}_2$ . *noise intensity* and *LL* are represented by  $\times$  and  $\bullet$ , respectively. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

Figure 30: Rolling estimates of the 2-type  $\beta$  model for FTSE

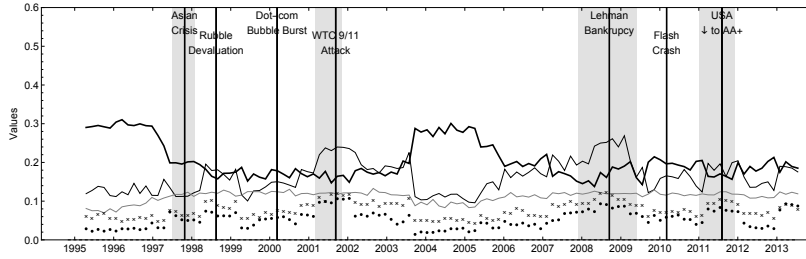
(a) MA61 fundamental price approximation



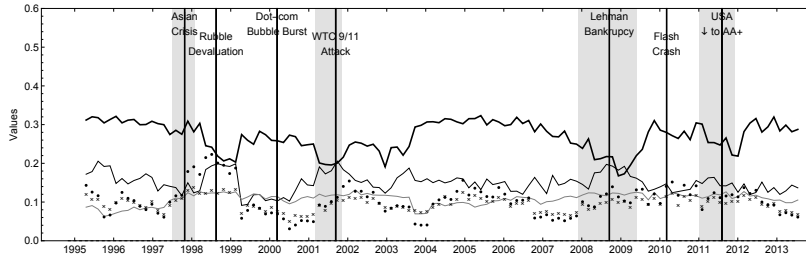
(b) MA241 fundamental price approximation



(c) MA61—related standard deviations



(d) MA241—related standard deviations

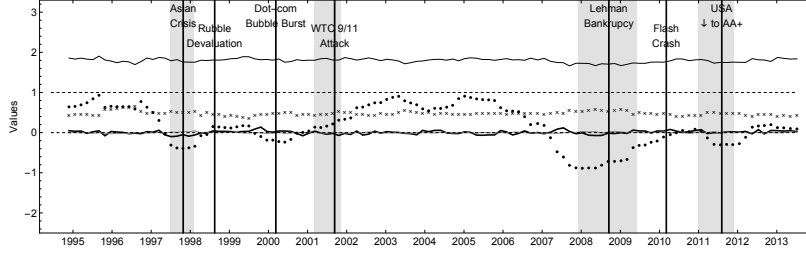


*Note:* Bold black full line depicts  $\hat{\beta}$ , black full line depicts  $\hat{g}_2$ , and grey full line depicts  $\hat{b}_2$ . *noise intensity* and *LL* are represented by  $\times$  and  $\bullet$ , respectively. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

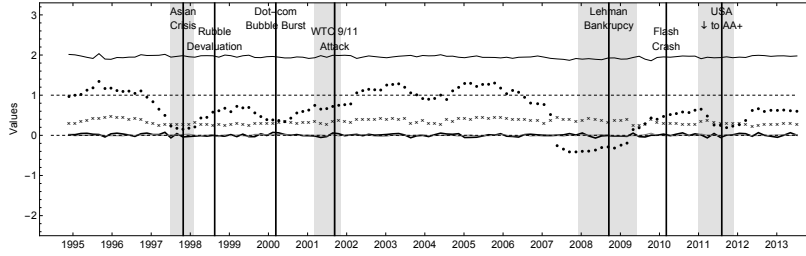


Figure 31: Rolling estimates of the 2-type  $\beta$  model for HSI

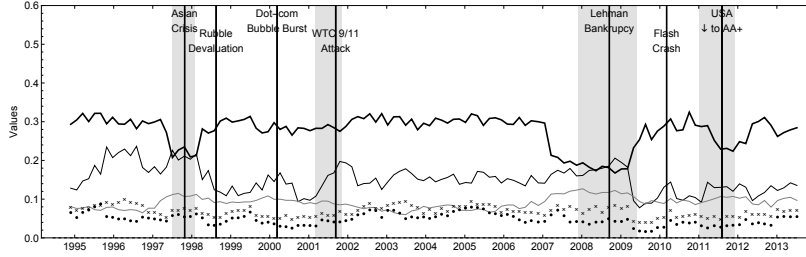
(a) MA61 fundamental price approximation



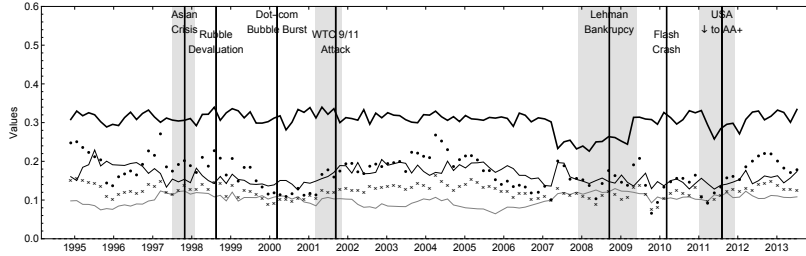
(b) MA241 fundamental price approximation



(c) MA61—related standard deviations



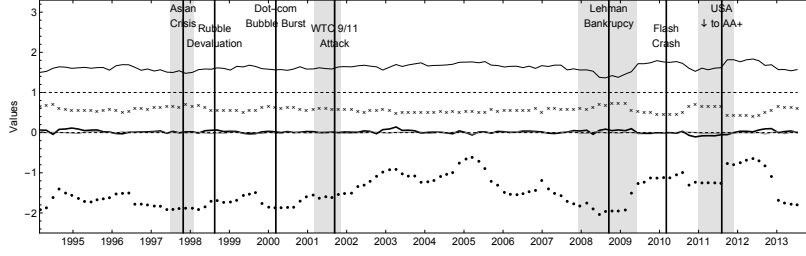
(d) MA241—related standard deviations



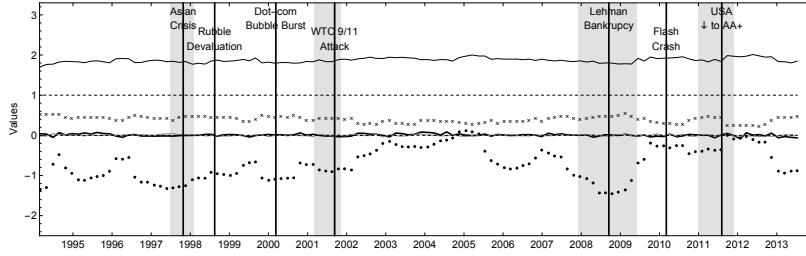
*Note:* Bold black full line depicts  $\hat{\beta}$ , black full line depicts  $\hat{g}_2$ , and grey full line depicts  $\hat{b}_2$ . *noise intensity* and *LL* are represented by  $\times$  and  $\bullet$ , respectively. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

Figure 32: Rolling estimates of the 2-type  $\beta$  model for NIKKEI 225

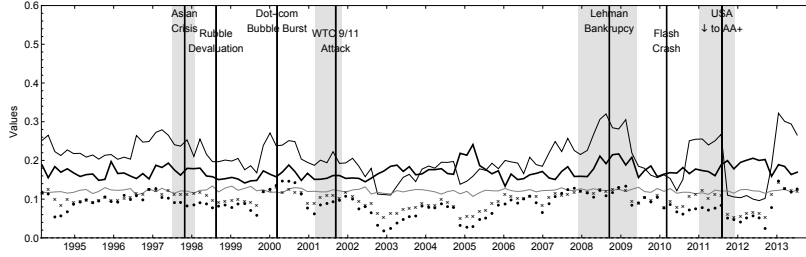
(a) MA61 fundamental price approximation



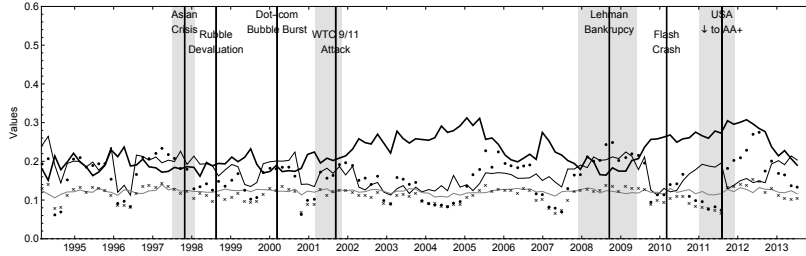
(b) MA241 fundamental price approximation



(c) MA61—related standard deviations



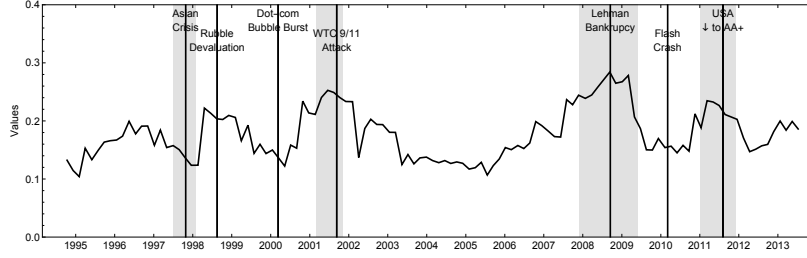
(d) MA241—related standard deviations



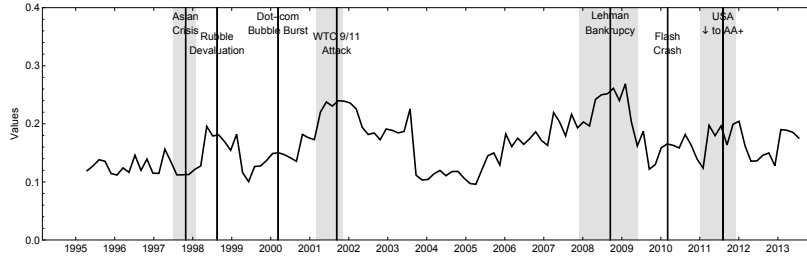
*Note:* Bold black full line depicts  $\hat{\beta}$ , black full line depicts  $\hat{g}_2$ , and grey full line depicts  $\hat{b}_2$ . *noise intensity* and *LL* are represented by  $\times$  and  $\bullet$ , respectively. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

Figure 33: Rolling behaviour of the SD of the  $\hat{g}_2$  estimate II.

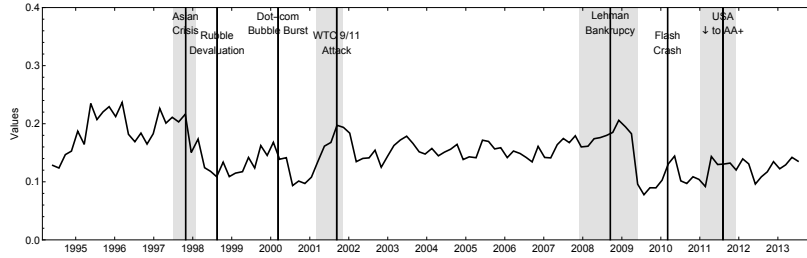
(a) DAX MA61 fundamental price approximation



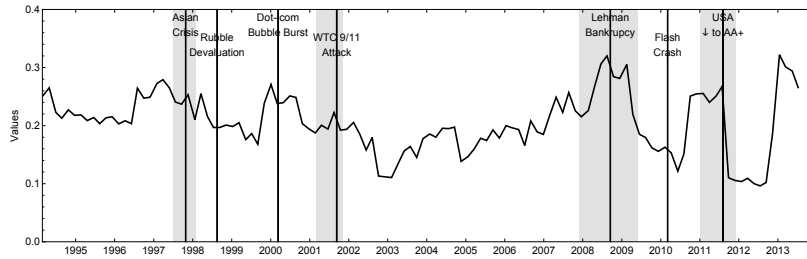
(b) FTSE MA61 fundamental price approximation



(c) HSI MA61 fundamental price approximation



(d) NIKKEI 225 MA61 fundamental price approximation



*Note:* Bold black full line depicts standard deviation of the  $\hat{g}_2$  estimate. Results are based on 200 random runs, length of the rolling window is 240 days with 40 days steps, and the kernel estimation precision  $N = 500$  i.i.d. draws from normal distribution. Sample medians are reported. The figure also depicts several important stock market crashes (marked as vertical black lines) and recession periods (depicted in grey).

# IES Working Paper Series

## 2016

1. Jiri Skuhrovec, Jan Soudek: *zIndex – Benchmarking Municipalities in Public Procurement*
2. Diana Zigraiova: *Management Board Composition of Banking Institutions and Bank Risk-Taking: The Case of the Czech Republic*
3. Tomas Havranek, Roman Horvath, Ayaz Zeynalov: *Natural Resources and Economic Growth: A Meta-Analysis*
4. Roman Horvath, Lorant Kaszab: *Equity Premium and Monetary Policy in a Model with Limited Asset Market Participation*
5. Jana Votapkova, Pavlina Zilova: *Determinants of Generic Substitution in the Czech Republic*
6. Vojtech Korbilius, Michal Paulus, Tomas Troch: *Life Expectancy and its Determinants in the Czech Republic*
7. Jiri Kukacka, Jozef Barunik: *Simulated ML Estimation of Financial Agent-Based Models*

All papers can be downloaded at: <http://ies.fsv.cuni.cz>



Univerzita Karlova v Praze, Fakulta sociálních věd

Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26

E-mail : [ies@fsv.cuni.cz](mailto:ies@fsv.cuni.cz)

<http://ies.fsv.cuni.cz>