MAGKS

Joint Discussion Paper
Series in Economics

by the Universities of
Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 28-2015

Alexandru Mandes and Peter Winker

Complexity and Model Comparison in Agent Based Modeling of Financial Markets

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Complexity and Model Comparison in Agent Based Modeling of Financial Markets†

Alexandru Mandes*
University of Giessen
Licher Str. 64
35394 Giessen, Germany
email: Alexandru.Mandes@wirtschaft.uni-giessen.de

Peter Winker
University of Giessen
Licher Str. 64
35394 Giessen, Germany
email: Peter.Winker@wirtschaft.uni-giessen.de

August 14, 2015

Abstract. Agent based models of financial markets follow different approaches and might be categorized according to major building blocks used. Such building blocks include agent design, agent evolution, and the price finding mechanism. The performance of agent based models in matching key features of real market processes depends on how these building blocks are selected and combined. For model comparison, both, measures of model fit and model complexity are required. Some suggestions are made on how to measure complexity of agent based models. An application for the foreign exchange market illustrates the potential of this approach.

Key Words: Agent based modeling, model selection, complexity

JEL classification: C63, C18, C58, G17

†We are indebted to Manfred Gilli and participants of the Second Meeting of the German Network for New Economic Dynamics (GENED) at the Technical University Darmstadt, 29th - 30th September 2014, for valuable comments on a preliminary version of this paper.

*The author would also like to thank Deutscher Akademischer Austauschdienst (DAAD) for the awarded PhD scholarship.

1 Introduction

Agent-based modeling (ABM[1]) basically stands for bottom-up modeling, i.e. instead of directly modeling the global outcome, the focus is shifted towards the individual agents and on the interactions between them, which subsequently lead to the emergence of collective behaviour. By this description of ABM, one can see that ABM might also be interpreted as a computational model suitable for describing complex systems. In the field of economics, this modeling framework is also known as Agent-based Computational Economics or Complexity Economics (see, e.g., Tesfatsion (2001), Arthur (2014)).

ABM has been applied in various fields of research where complex systems are involved, ranging from natural phenomena (e.g., ecological systems as in Reuter, Breckling and Jopp (2011)) to socio-economical systems (see, e.g., the contributions in LeBaron and Winker (2008) for some examples). This contribution focuses on applications to financial markets, but the approach taken might be useful also in other fields.

Reviewing the literature on ABM in the financial markets context allows to identify two slightly different lines of research. The first one, which covers most of the existing studies, views ABM as a tool for testing abstract ideas or thought experiments by means of computer-based simulations. Such ABM are usually evaluated only qualitatively, if at all, e.g., by examining their ability to replicate some of the so-called "stylized facts" of financial markets (Cont, 2001; Winker and Jeleskovic, 2006; Winker and Jeleskovic, 2007). The second line takes a more quantitative (econometrical) approach and tackles calibration to real data (Winker, Gilli and Jeleskovic, 2007) or even forecasting.

A calibration procedure – in the sense of parameter estimation – represents a relevant step in order to fit a model to a specific time series. However, given the uncertainty about the most adequate model, choosing parameter values alone is not sufficient. It has to interact with a further step, namely taking into account also the model specification, i.e. a model selection procedure. A first example is provided by Winker et al. (2007) who rejected the model proposed by Kirman (1991, 1993) for exchange rate data under the assumption of a constant fundamental value. The rejection was based on an objective function comparing properties of real and simulated data. The minimum obtained for this objective function was tested against the bootstrap distribution for the real data. Winker et al. (2007) also rejected the model suggested by Lux and Marchesi (2000). A further example is provided by Fabretti (2013) who set up an optimization problem for the ABM proposed by Farmer and Joshi (2002) using the objective function introduced by Winker et al. (2007) and, both a Nelder-Mead with Threshold Accepting, as well as a Genetic Algorithm heuristic. After fitting the model parameters and comparing the (optimized) simulated time series properties with the ones of real data, the authors conclude that the replication is not completely

[1] We will also use the acronym ABM for agent based models.

satisfactory, which is attributed to the design of the model itself.

Model design is still quite an open problem in ABM. In the last two decades, a substantial number of ABM for financial markets have been proposed following quite different approaches. This diversity might be explained by the interdisciplinary character of ABM and, consequently, researchers with a different modeling background. Chen (2012) provides a historical perspective on the different lines of research in agent-based computational economics. Moreover, several surveys review the most prominent milestones of this development, e.g., LeBaron (2000), Tesfatsion (2001), Hommes (2006), LeBaron (2006), Lux (2009), Chiarella, Dieci and He (2009), Hommes and Wagener (2009), Cristelli, Pietronero and Zaccaria (2011), Iori and Porter (2012).

However, a commonly agreed upon set of "core concepts" is still missing, and also the issue of discriminating between various models has not been completely solved. Chen, Chang and Du (2012) propose an ABM taxonomy and order the different types of models with respect to their complexity.[2] According to the authors, there are essentially two main design classes of financial agents in ABM of financial markets: the N-type design and the autonomous-agent (AA) design.

Within the first class, agents have either a behavioral forecasting rule, such as fundamentalists and chartists, or follow a more loosely-typed rule described by general – linear or non-linear – functions. The adaption strategy consists in a herding or a probabilistic switching mechanism. This class spans from few-type to many-type agents. Intuitively, the few-type models are considered less complex than the many-type models because of their limited heterogeneity.

On the other side, the autonomous-adaptive agents are associated with a more complex learning algorithm, such as a Genetic Algorithm (GA) or Genetic Programming (GP), which does not only change the agents' behavior by choosing between predetermined strategies, but also creates new, sometimes improved, trading rules. Again intuitively, this class is considered to be more complex than the N-type design, as the learning mechanism is able to generate a more heterogeneous population, which enables to search through a larger space of forecasting functions.

Chen et al. (2012) classify 50 ABM with respect to their type (class membership: 2-type, 3-type, many-type, AA) and analyzes to what extent the type has an impact on the capability to replicate stylized facts. The authors find no significant evidence in favor of more complex design models having higher explanatory power (explaining more stylized facts).[3] From our perspective, it is interesting that Chen et al. (2012) define the concept of complexity for an ABM class as the degree of agents' heterogeneity (diversity).

[2]It should be noted that the term "complexity" is used here in a quite general meaning. We come back to the question which concepts and measures of complexity might be adequate in the context of ABM in Section 3.

[3]It is also to be noted that the results are based on original reported data and not every ABM approach refers to the same list of stylized facts.

Though, this heterogeneity is not independent of the learning algorithm.

Following a different approach, Westerhoff (2009) considers the “level” of building blocks of ABM in order to find explanations for their ability to replicate stylized facts. The author tweaks around with different parametric forms and with stochastic factors of the agents’ trading rules, of the strategy selection mechanism, as well as of the price impact function, which defines the price adjustment process. He finds that the underlying structure of the endogenous dynamics is a result of the non-linear (deterministic or stochastic) interactions between positive and negative feedback rules.

Our research question is how size and complexity of an ABM can be defined at the conceptual level, and how it actually could be measured. Thereby, we have to take into account that from the perspective of a complex system, at least two dimensions have to be distinguished: The size of the model itself (number of components, parameters, description length, non-linearity etc.) and the emergent complexity resulting from the aggregation of individual behavior in an ABM. We might consider the first approach as the more conventional input oriented definition of complexity, while the second one is focusing more on the qualitative properties of the resulting dynamics.

We will start by identifying the main building blocks of typical ABM for financial markets in Section 2. In the following Section 3, we will discuss some facets of complexity and how they could be accommodated with ABM. The approach is illustrated with an example of the foreign exchange market in Section 4. Section 5 concludes and provides an outlook to further research.

2 Building Blocks of Agent Based Models of Financial Markets

Generally, an ABM consists of interacting adaptive agents. In the case of ABM for financial markets, the agents play the role of traders and their interactions might be direct, implemented by various social learning processes, or indirect, represented by individual evolution strategies centred around the endogenous market price (trading demand interaction). In this section, we will describe in more details the three main building blocks, i.e. the agent itself, the evolution/learning and the price discovery mechanisms.

2.1 Agent design

The agent is equipped with one or several trading strategies, which transform information input into conditional expectations about the future price or return. Actually, deriving beliefs is the main component of the agent’s bounded rationality, and the heterogeneity of agents’ expectations is one of the pillars of ABM. In a second step, the investment decision, i.e. the trading demand is derived by taking into consideration risk preferences or portfolio selection models. In the case of intraday ABM, when individual trades are modeled, it is also important how agents actually trade, i.e. the assumptions about order execution.

The simplest and maybe the earliest trading strategy used in ABM is the random behavior or zero-intelligence (ZI) strategy. ZI agents can be generalized as in Gode and Sunder (1993) to ZI-constrained (ZIC) traders, which are subject to some constraints, e.g., resource or bankruptcy restrictions, or as in Cliff and Bruten (1997) to ZI-plus (ZIP), i.e. profit motivated traders. The class of ZI strategies represents an important reference to which more complex ABM can be benchmarked and is also used in order to evaluate other ABM components such as the institutional or market design (price discovery).

More structure is added by means of parametric forecasting rules. These mathematical models vary in shape as well as in size, e.g., with regard to the number of parameters and information variables.[4] Some functions correspond to stylized implementations of rule-governed trading strategies, which mimic human behavior. The standard chartist-fundamentalist dichotomy, backed by empirical evidence provided in survey studies, e.g., by Frankel and Froot (1987), was implemented since the early developments of ABM, e.g., Kirman (1991, 1993). Fundamentalists – also called "value investors" – build their expectations about future prices upon assets' fundamentals, which are cumulatively reflected by a fundamental value. They believe that, in the long run, the market price should converge to this intrinsic value. Consequently, they buy undervalued and sell overvalued assets.[5] On the other side, technical traders create trading rules that extrapolate observed price patterns. A particular kind of technical traders are the trend followers, who believe that price changes have inertia also known as momentum. This kind of strategy adds a positive feedback to the dynamics of financial markets, and may be destabilizing. Another type of chartists known as contrarians, rely on reversal price formations and thus contribute a negative feedback on the market price.

Within the type of non-parametric forecasting rule, we can mention the condition classifier proposed by Palmer, Arthur, Holland, LeBaron and Tayler (1994), Arthur, Holland, LeBaron, Palmer and Tayler (1997), and LeBaron, Arthur and Palmer (1999), which is a non-continuous, piece-wise linear forecasting rule. Also, Chen and Yeh (2001) have implemented agents based on genetic programming trees (GPT). The authors are thus able to define the complexity of the forecasting model based on the number of nodes or based on the depth of the tree. These designs are strongly related to evolutionary algorithms – genetic algorithms (GA) and genetic programming (GP) – ap-

[4]Usually, one of the factors is stochastic in order to reflect missing information or uncertainty about the strategy. It also introduces some heterogeneity of agents' behavior.

[5]The fundamental value can be perceived by the agents without or with errors. For example, in Fischer and Riedler (2014), evaluation errors persist only for a limited time, as agents will eventually become aware and correct them.

plied in order to adapt and evolve agents' trading strategies, as discussed in the following Subsection 2.2.

With respect to their risk awareness, agents can be assumed to be either risk neutral (only aiming at maximizing their profits), or risk averse, i.e. determining their optimal portfolio composition by trading-off expected return against expected risk. In the later case, the asset demand function is usually computed in the framework of expected utility maximisation. Under constant absolute risk aversion (CARA), the optimal position in the risky asset the agent wishes to hold is independent of wealth. On the other side, the constant relative risk aversion (CRRA) computes the proportion of total wealth the agent wants to invest in the risky asset in the up-coming period. Thus, in this setting, the demand for the risky assets will increase with wealth in a linear way.

Most ABM ignore the agents' individual wealth in computing the asset demand and rely mainly on the expected return. By including the wealth component into the model, an agent's demand becomes restricted also by her current wealth, which might add to the heterogeneity of agents in the model and also introduces an extra feedback loop – when balance sheets are marked to market, prices directly affect a trader's wealth, which in turn has an impact on the trader's ability to influence future prices (see, e.g., Fischer and Riedler (2014)).

2.2 Evolution and Learning

It is possible to differentiate between two main approaches for strategy selection, which correspond to the two directions of building ABM identified by Chen et al. (2012). In the first approach, which might be labeled as *herding*, a switching mechanism is implemented, which allows agents over time to change their type out of a set of two or more (N-type) options corresponding to alternative strategies. The second approach consists in inductive learning, which is related to computational models of autonomous agents.

Within the group of switching mechanisms, a further distinction can be made between two different types of switching methods. The first one, i.e. probabilistic switching, is purely stochastic and describes herding behavior by recruitment or opinion formation driven by the social pressure of the majority. For example, Kirman (1991, 1993), proposes a mechanism where agents meet at random and the first is converted to the second one's view with a given probability; there also exists a small probability that the first agent will change his opinion independently. The second type is represented by the case of evolutionary switching. In this setting, the switching probabilities which determine whether the agent sticks to or changes her strategy are not constant any more, but depend on the recent relative performance of the strategies. In Brock and Hommes (1997, 1998), the evolutionary selection of trading strategies is given by the multinomial logit probabilities from a discrete choice model. Under the switching paradigm, agents are not fully heterogeneous as they randomly choose from the same distribution of potential strategies. In other words, Chen et al. (2012) point out that, under this paradigm, "all agents are homogeneous *ex ante*, but heterogeneous *ex post*".

For the second approach, random initialized strategies are updated by means of an evolutionary algorithm. Individual learning might be considered to be a more realistic type of evolution, as agents' strategies are typically not observable in reality and, thus, direct interaction and imitation between different agents cannot always be assumed. Instead, agents can only observe the final outcome, i.e. the price endogenously driven by all of the agents' expectations and their resulting decisions,[6] and learn only from their own experience. A typical implementation of inductive learning is given by the Santa Fe Institute Artificial Stock Market (SFI-ASM), described in Palmer et al. (1994), Arthur et al. (1997), and LeBaron et al. (1999). In this model, a multi-population GA is run within each agent.[7] A different approach which implements a single-population GP is proposed by Chen and Yeh (2001).[8] Alternatively to the concept of individual learning, social learning assumes direct interaction between the agents and is usually implemented by applying a single-population GA/GP directly over the entire population of market participants.

The search incentive, i.e. the agent's decision to change her current strategy, can also be modeled in an explicit way. In Chen and Yeh (2002) it is the result of a two-stage independent Bernoulli experiment, with the success/failure probabilities depending on the peer pressure (a rank of traders' performance) and on the self-pressure (individual progress over the last period of time). Alternatively, Fischer and Riedler (2014) model the probability that the agent sticks to her current strategy with a binary-choice model, by comparing a measure of individual profit to a benchmark representing the average of all agents' profitability measures.

2.3 Price discovery

Traditional ABM focus on trading at a low (daily) frequency and the price discovery mechanism resemble call auctions, where trading orders are batched together and executed at the same price and at the same time.[9] The new

[6] Arthur et al. (1997) underline this reflexive nature of the market, where agents' expectations co-evolve in a world they co-create.

[7] Each agent has at her disposal a set of strategies which is continuously evaluated – even if the rule was not actually executed, but would have been a valid selection based on its condition classifier part – and updated at random times by replacing a subset of the agent's worst rules with new ones generated by crossing-over and mutating a selection of its own best rules.

[8] A population of new trading rules is evolved by means of GP within a distinct "schooling" component, which is visited from time to time by underperforming single strategy holding agents.

[9] Alternatively, in the case of intraday ABM, a continuous market is implemented where trading demand is disclosed asynchronously and orders are matched by means of a limit

market equilibrium price is updated usually through a market impact function (e.g., Lux (1995, 1998), Lux and Marchesi (1999, 2000), Chen and Yeh (2001, 2002), Farmer and Joshi (2002), Westerhoff (2009, 2010)) or by Walrasian tâtonnement (e.g., Arthur et al. (1997), Brock and Hommes (1997), Fischer and Riedler (2014)). In the former case, the price change is a function of the aggregate excess demand. This function varies in shape and can be interpreted as the speed of price adjustment. The excess demand is assumed to be cleared by the market maker at the new adjusted price. The latter, i.e. the Walrasian tâtonnement procedure, identifies by means of numerical analysis an equilibrium where there is no excess demand and all agents trade at this clearing price.

3 Complexity of Agent Based Models

When comparing alternative models, a standard approach consists in considering two aspects. First, one analyzes to what extent the model is able to describe those features of the analyzed objects the researcher is most interested in. In an ABM context this might be idealized properties, e.g., endogenous creation of bubbles, or features related to real data, e.g., generating stylized facts or even well fitting – in an econometric sense – selected statistical moments of the data. Second, the principle of parsimony suggests to prefer a "small" model to a larger one if both exhibit the same degree of fit. In a typical linear econometric model, the "size" of a model can be described by the number of parameters to be estimated. Then, information criteria such as the ones suggested by Akaike (1974), Schwarz et al. (1978) and Hannan and Quinn (1979) are well established tools for analyzing the trade-off between good description and parsimony (Haughton, 1991).

It turns out that in the framework of ABM for both aspects of model evaluation, i.e. quality of description and complexity, there do not seem to exist concepts which are generally agreed upon let alone broadly applied. Gilli and Winker (2003) and Winker et al. (2007) introduced the idea of simulated method of moments for the evaluation of ABM in applications to financial markets. Although it is still an open issue which moments should be taken into account, this and similar approaches are used more routinely over the last few years (see, e.g., Franke (2009), Fabretti (2013), Grazzini and Richiardi (2015)). While this might provide an answer to the first issue, namely measuring descriptive quality or fit, a standardized answer to the second question, namely on how to measure the complexity of an agent based model, seems to be still far off. We do not claim to provide a conclusive answer to this issue. Rather do we try to pave the way for a discussion of

order book at various prices (see, e.g., Chiarella and Iori (2002), Daniel (2006), Chiarella, Iori and Perelló (2009), Mandes (Forthcoming)). As a consequence, transactions do not take place only at the global equilibrium price. However, this higher-frequency framework deals with a finer level of details which are not within the scope of market features we want to capture in this paper.

what "complexity" could mean for an ABM, which eventually might result in an accord on how to measure complexity of an ABM.

In the case of ABM, or more generally of complex systems, the topic is not trivial given that there does not exist any unified definition, let alone standardized measures of complexity as concepts of complexity depend on the specific context. Lloyd (2001) has collected a list of concepts and measures of complexity used in quite diverse settings. He identifies three different dimensions that researchers have used in order to quantify complexity: difficulty of description, difficulty of creation, and degree of organization. In the following subsections, we try to organize our discussion of complexity of ABM along these dimensions.

3.1 Difficulty of Description

Difficulty of description refers to the model. In the setting of a linear regression model, such a measure would be the number of parameters. Obviously, the model

$$y_t = \alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \varepsilon_t \tag{1}$$

is more difficult to describe than

$$y_t = \alpha_0 + \alpha_1 x_{1,t} + \varepsilon_t \,. \tag{2}$$

However, considering a further, non-linear model

$$y_t = \alpha_0 + \alpha_1 x_{1,t}^{\alpha_2} + \varepsilon_t \,, \tag{3}$$

the ranking, e.g., between model (1) and (3), does not seem to be as obvious any more. Nevertheless, the use of the number of parameters to be estimated is common in the context of information criteria also for non-linear settings (Nakamura, Judd, Mees and Small, 2006). Therefore, we will also include this standard measure of difficulty of description as one option for describing ABM complexity keeping in mind its limitations with regard to highly non-linear processes. We will label this measure by $\mathcal{C}_{\#P}$, complexity according to the number of parameters.

A related measure of difficulty of description could take into account the effort needed to describe the actual proceeding of an ABM. When making use of a standard description language, e.g., based on ideas of a Turing machine, the length of the shortest algorithm describing the model might be a sensible measure (for an introduction in standard computational complexity see, e.g., Du and Ko (2014)). Based on this measure, model (3) would exhibit a substantially higher complexity compared to model (1). Unfortunately, deriving the complexity of an ABM in this sense, is a highly complex task by its own. First, typically, ABM are implemented in some higher level language. Thus, statements in this language would have been to recorded in the standard description language also taking into account components used from external libraries, e.g., random number generators. Second, even the length of such a machine level code would represent only an upper bound for the actual complexity. In addition, a lower bound is required, which is an even more demanding challenge.

Therefore, we will rely in our analysis on a crude method, which is only an approximation to an upper bound, namely the length of the code used for the ABM. When using higher order programming languages, the comparison has to be based on the same language for alternative models, ideally using similar conventions for naming of variables and functions, inclusion of external libraries etc. Furthermore, comments in the code as well as parts of the code dealing with result presentation etc. have to be removed prior to obtaining the approximation. Despite of these limitations, we will also consider this crude approximation of relative difficulty of description and denote it by $\mathcal{C}_{code}$.

Alternatively, software complexity metrics, created for development and testing purposes, could be used. The cyclomatic complexity number is an example of such a quantitative measure. It is based on the number of linearly independent execution paths through a program's source code, i.e. each function or method has a complexity of one plus one for each branch statement such as *if*, *else*, *for*, *foreach*, *while*, *case* in a switch block or *catch* or *except* statement in a try block (for details see McCabe (1976), Watson, McCabe and Wallace (1996), Rosenberg (1998)).[10] When a project has multiple source files, both a maximum and an average cyclomatic complexity metric can be reported, further denoted by $\mathcal{C}_{cyc-max}$ and $\mathcal{C}_{cyc-av}$. Similarly, the maximum block depth, i.e. the maximum nested block depth level found,[11] and the average block depth, i.e. the average nested block depth weighted by depth, can be computed – further labeled by $\mathcal{C}_{depth-max}$ and $\mathcal{C}_{depth-av}$.

3.2 Difficulty of Creation

Difficulty of creation measures what it takes to generate the outcome of a model. In the context of ABM for financial markets, such a measure could be the computational time needed to generate, e.g., an aggregate price time series of a given frequency and length. Obviously, in order to make such a concept operational for the comparison of several models, one has to run all of them in an identical setting (software, operating system, hardware).[12] Furthermore, run time might depend on parameter values, seeds of random

[10] A complexity count is added for each logical connective operator *and* and *or* within logical statements. Also, conditional expressions using the ternary operator `?:` add one count to the complexity total.

[11] At the start of each file the block level is zero and the block depth grows with each execution control statements such as *if*, *case* and *while*.

[12] One way to achieve such a standardization consists in considering Flops instead of time. Given that this might not be feasible in practice, counting the counts to procedures, functions etc. might provide a way to obtain more reliable lower bounds for this type of computational complexity.

number generators or the status of other routine operations of the computer in the background. In particular, for some parameter settings, the behavior of the model might become more chaotic resulting in a higher computational load. Therefore, the evaluation of run time should be performed either for a set of relevant parameter values and several random seeds or for those parameter values which are the result of an optimization step.

Finally, we have to take into account that this empirical measure of run-time for a specific model provides only an upper bound to the actual difficulty of creation, as the model might be implemented in a more efficient way reducing computational time substantially. Given that obtaining an accurate measure of the minimum computational time required to generate a model's output is a difficult task similar to the one of algorithmic complexity described in the previous subsection, we will have to stick to a rather heuristic approximation of the upper bound as provided by timing tools. We will denote this measure of complexity by $\mathcal{C}_{time}$. Similarly, a measure regarding the memory storage requirements can be defined, henceforth labeled $\mathcal{C}_{space}$.

3.3 Degree of Organization

A potential downside of the measures introduced so far is that they focus only on static (quantitative) dimensions of an ABM. It might be challenged whether these are sufficient to account for the non-linear interactions which lead to unpredictable collective behavior or for the continuous evolution of agents due to the learning process. In other words, two ABM of equal size, e.g., number of parameters, might be capable of generating different potential outcomes because of the way how the building blocks are connected. This aspect is addressed by measures of the degree of organization, which puts emphasis on the amount of exchange between the components of the model. For an ABM, e.g., by design only, the degree of organization might be considered to be higher, if there is direct interaction between agents than if interaction takes place exclusively over the market price. It becomes obvious, that all these dimensions focus more on the architecture of the model, i.e. its structural relations, than on its output, i.e. the type of (global) dynamics generated by the model.

Since many ABM implement a learning mechanism (as described in Subsection 2.2), we could also look at the emergence of self-organisation which, alongside adaptiveness, is a key feature of complex systems. This layer can be considered as being intermediary between the static blueprint of the model and its final output, e.g., the price time series. From an organic point of view, an ABM can be seen as an information processing mechanism with the final goal of providing efficiency – a market price in line with the fundamentals – under bounded rationality, as well as trading liquidity. In its dynamics, new information is being handled at every time step. Besides the obvious exogenous shocks, e.g., changes in the fundamental price, one can also identify an internal (endogenous) source of information represented by the current agents' wealth and strategies. These structures actually act as a memory component, which is storing information describing the current state of the market, is being updated at each simulation step and constitutes a prior for the next steps. We argue that how these internal distributions evolve could be considered as a more qualitative trace of the complexity of ABM, reflecting also the latent features of the system, such as feedback loops, which are difficult to be identified and measured otherwise.

On one side, a higher degree of agents heterogeneity – captured quantitatively by $\mathcal{C}_{space}$ – corresponds to a bigger storage capacity,[13] but this is not a sufficient condition to reckon a more "capable" system, as the latent information should also be fructified in a meaningful way. The patterns of how these middle-layer distributions change are key. One can distinguish between two extremes, i.e. regularity (periodicity) and randomness, both corresponding to simple behavior. According to Mitchell (2009), "complexity" is related to a combination between order and randomness. More precisely, a complex entity can neither be too ordered, nor too random, and its effective complexity is represented by the amount of information contained in its regularities.

The approximate entropy $ApEn(m, r)$, defined in Pincus, Gladstone and Ehrenkranz (1991) and Pincus (1991), is a model-independent statistic which quantifies the amount of sequential regularity and can potentially discriminate between classes of systems, e.g., periodic, deterministic chaotic and stochastic. The two parameters which need to be specified stand for the block length m and the tolerance window r.[14] The $ApEn$ algorithm computes a nonnegative number, with smaller values corresponding to more regular or ordered sequences and larger values assigned to more irregular or random time-series. $ApEn$ can be used as a relative measure which can distinguish between different groups of data for fixed m and r. If the groups have significantly different standard deviations, a normalized $ApEn$ is recommended for comparison purposes. A shortcoming with respect to ABM is that the approximate entropy can only be applied to univariate time-series, while in the case of more than two agents we are dealing with multivariate data. Therefore, we will use it only for 2-type models, such as the Kirman, Lux and Farmer-Joshi models (see Section 4), and label it $\mathcal{C}_{ApEn}$. If applied directly to the final (observable) output, this statistic could evaluate and even validate/reject a calibrated model with respect to specific data.[15]

Similarly, the patterns of time-series describing the dynamics of the endogenous information contained in the middle-layer could also be analysed

[13] This view is in line with the intuition of Chen et al. (2012) who define complexity as the degree of heterogeneity (diversity).

[14] The implementation algorithms are described in Pincus et al. (1991) and Manis (2008).

[15] There is also the potential of becoming a stylised fact, given that certain values would be exhibited by a wide range of financial data. A first application is contained in Pincus and Kalman (2004), but more work is needed. As benchmark, Pincus and Kalman (2004) have computed the $ApEn(1, 0.2\,\sigma)$ for the DJIA log returns at a 10-minute time frequency and obtained a maximum value of 1.949, with all moving window values oscillating in a range above 80% of the maximum.

based on their fractal dimension index, which gives information about the geometrical structure of the system at multiple time scales (for an introduction to fractals and the concept of fractal dimension see, e.g., Mandelbrot (1977) and Mitchell (2009)). The fractal dimension quantifies the complexity of a system as a ratio of the change in detail to the change in scale. A technique for calculating the fractal dimension of a time-series was proposed by Higuchi (1988), where the estimate of the fractal dimension – labeled in the following by $\mathcal{C}_{HFD}$ – is given by the exponent of the power law distribution of the average lengths of curves extracted at different scales (the algorithm is briefly described in Appendix A).

Another complexity measure known as statistical complexity has been proposed by Crutchfield and Young (1989). This measure is non-monotonic in the sense that it is small for simple systems (highly ordered and random ones) and high for the in-between systems, which are intuitively considered to be complex. The analysed information needs also to be first translated into a sequence via a measuring channel. Thus, the ABM can be considered a "message" source and the general idea is to construct a "parallel" computational model which is able to "predict" its behavior (statistically indistinguishable). The information content of the simplest such model determines the complexity class of the original system. Alternatively, the Kullback-Leibler (KL) distance between the simulated time-series and a random walk benchmark could be used as a measure of model complexity. Barde (2015) develops an Universal Information Criterion (UIC), which relies on mapping the target data-generated process to an underlying Markov process by applying a universal data compression algorithm, i.e. the Context Tree Weighting (CTW) algorithm proposed by Willems, Shtarkov and Tjalkens (1995). Once the Markov transition matrix is generated, the mean value of the benchmark observation-level score, i.e. the log-likelihood, can be computed, giving the prediction accuracy which is equivalent to the KL distance. Because computing these measures is quite tedious, we will not apply them at the current step of our research.

4 An Application to the FX Market

We first compute the measures of complexity described in Section 3 for three ABM of the behavioral class – the *Kirman model* introduced in Kirman (1991, 1993) with a detailed version specified in Gilli and Winker (2003), the *Lux model* described in Lux (1995, 1998) and Lux and Marchesi (1999, 2000), and the *Farmer-Joshi model* from Farmer and Joshi (2002).

4.1 Models specifications

The Kirman model

The Kirman model is a 2-type design model, where N agents are split between fundamentalists and chartists. At every step,[16] any agent $i = 1 \dots N$ can switch his type by random mutation with probability ε or by direct interaction with another agent with probability δ. Thus, we can classify this procedure as probabilistic switching. However, the final trading decision does not depend directly on the agent's type, but on a majority assessment procedure, where each agent is able to observe the share of fundamentalists in the market $q_t = \#\{i|i \in \mathcal{F}\}/N$ under the form of a noisy signal $\tilde{q}_{i,t} \sim \mathcal{N}(q_t, \sigma_q^2)$. The agent's expectations about the future log price change $E_t^i[\Delta p_{t+1}]$ will be in line with the perceived majority, i.e. either linear fundamentalist (see equation (4)) or linear chartist (see equation (5)).

$$E^i_{F,t}[\Delta p_{t+1}] = \alpha_F\,(\bar{f} - p_t), \forall i \in \mathcal{F}, \tag{4}$$

where $0 < \alpha_F \leq 1$ is the adjustment speed in fundamentalists expectation and $\bar{f}$ is the long-run equilibrium given by the mean of the logarithm of the fundamental value.

$$E^i_{C,t}[\Delta p_{t+1}] = (p_t - p_{t-1}), \forall i \in \mathcal{C} \tag{5}$$

Price updating takes the form of an Walrasian tâtonnement mechanism, where the equilibrium price $\hat{p}_t$ – at which there is no excess demand – results from solving the market model specified as follows:

$$p_t \stackrel{!}{=} c\,E_t^M[\Delta p_{t+1}] + \bar{f}, \tag{6}$$

where c is the market maker's price adjustment sensitivity and $E_t^M[\Delta p_{t+1}] = \sum\limits_{i=1}^{N} E_t^i[\Delta p_{t+1}]/N$ is the aggregate market expectation.

An analytical solution of (6) is possible and is given by:

$$\hat{p}_t = \frac{(1 + c\,w_t\,\alpha_F)\,\bar{f} - c\,(1 - w_t)\,p_{t-1}}{1 + c\,w_t\,\alpha_F - c + c\,w_t}, \tag{7}$$

where the only input variable is the fraction of agents "acting" as fundamentalists $w_t = \#\{i|\tilde{q}_{i,t} \geq 1/2\}/N$.

[16] Each trading day is divided into $n = 50$ micro-time intervals.

Finally, the next log price $p_{t+1} = \hat{p}_t + u_t$ is obtained by adding an exogenous perturbation $u_t \sim \mathcal{N}(0, \sigma_s^2)$, where σ_s^2 is the variance of fundamental shocks.

There are two possible implementations of the Kirman model. The first one consists in an holistic, explicit modeling of agents' actions and interactions as described previously. A second one is more compact and is based on the observation that the entire population can be conceptually reduced to only two clusters, since the parameters of the two strategies are fixed in time and common for all agents adopting the respective type of strategy.[17] The switching mechanism between the two clusters can be described by a stochastic process, where the dynamics of number of fundamentalists is given by:

$$
\begin{aligned}
P(i \to i+1) &= \left(1 - \frac{i}{N}\right) \left(\varepsilon + \delta \, \frac{i}{N-1}\right) \\
P(i \to i-1) &= \frac{i}{N} \left(\varepsilon + \delta \, \frac{N-i}{N-1}\right)
\end{aligned} \tag{8}
$$

and consequentially,

$$
P(i \to i) = 1 - P(i \to i+1) - P(i \to i-1).
$$

The aggregate price change expectation $E_t^M[\Delta p_{t+1}]$ can be viewed as a probabilistic average of the two cluster strategies, given the noisy perceived weight of fundamentalists in the market w_t, as defined in the following:

$$
E_t^M[\Delta p_{t+1}] = w_t \, E_{F,t}[\Delta p_{t+1}] + (1 - w_t) \, E_{C,t}[\Delta p_{t+1}] \tag{9}
$$

Winker et al. (2007) reformulate the probabilistic cluster weight w_t by applying, at each time step, the complementary normal cumulative distribution function as follows:[18]

$$
w_t = 1 - \Phi((0.5 - q_t)/\sigma_q^2), \tag{10}
$$

where Φ is the cumulative distribution function of the normal distribution $\mathcal{N}(0, 1)$.

As a bottom line, the second implementation requires less computational resources than the former, but cannot be applied to all agent based models, e.g., computational class ABM. In the case of the Kirman model, we will implement both and compare the results.

[17] Farmer and Joshi (2002) show that even if the strategy coefficients would be heterogenous among agents, when the strategies are linear, equivalent results can be achieved by using a "representative agent" for each strategy type with a coefficient equal to the group-mean. As a side note, Chen et al. (2012) classify this category as "few-type" models.

[18] Some "fallacy of composition" might arise due to this aggregation.

The Lux model

The Lux model is a 3-type design model based also on the fundamentalist-chartist feedback, but with three agent categories, i.e. fundamentalists N_F, optimistic N_C^+ and pessimistic chartists N_C^-.[19] The fundamentalist demand is given as a function of the mispricing from the fundamental value F (see equation (11)),[20] and the chartist demand is a function of the opinion index $x = (N_C^+ - N_C^-)/N_C$, reflecting the current market sentiment based on the relative number of optimistic and pessimistic chartists (see equation (12)).[21]

$$D^i_{F,t+1} = \alpha_F \, (F - P_t), \tag{11}$$

where $\alpha_F > 0$ is the reaction speed of fundamentalists.

$$D^i_{C,t+1} = x_t \, t^C, \tag{12}$$

where t^C is a fixed amount assets bought or sold by a chartist.

The switching mechanism between the three agent clusters is evolutionary, being dependent on the performance of the respective strategies. Chartists' switching between optimistic and pessimistic is based on the preference function U_1 which depends on the opinion index and the recent price change $\Delta P_t = P_t - P_{t-0.2}$ as in equation (13), where the time tick $\Delta t = 0.01$.[22] Thus, the change in the number of optimistic traders is $\Delta N_C^+ = N_C^- \, r^{-+} - N_C^+ \, r^{+-}$, where $r^{-+} \sim \mathcal{N}(\pi^{-+}, 1/N_C^-)$ and $r^{+-} \sim \mathcal{N}(\pi^{+-}, 1/N_C^+)$.[23]

$$\pi^{-+} = \frac{v_1}{n} \left(\frac{N_C}{N} \, e^{U_1} \right); \; \pi^{+-} = \frac{v_1}{n} \left(\frac{N_C}{N} \, e^{-U_1} \right); \; U_1 = \alpha_1 \, x + \alpha_2 \, \frac{\Delta P}{v_1}, \tag{13}$$

where π^{-+} denotes the directional probability for a pessimistic chartist to turn into an optimistic one, model parameter v_1 is the frequency of the respective transition type, $N_C/N = 1 - N_F/N$ stands for the fraction of

[19] Chen et al. (2012) identify the Lux model as an hierarchical two-type model, with two chartist subdivisions.

[20] The fundamental value is assumed to remain constant over the entire simulation time span.

[21] Agents' portfolios are ignored and therefore they are able to accumulate unbounded inventories.

[22] This means that there are $n = 100$ microintervals per day, at which interaction and trading sessions occur.

[23] When computing the effective net transitions between different clusters, we are not using the expected value as described in Lux (1998), rather we simulate the sampling data effect by drawing a random deviate from a normal distribution centred around the expected value and with a variance depending on the cluster size. Otherwise, due to rounding and very small probabilities per time unit, the number of agents changing their types would be zero most of the time.

agents not involved in the fundamentalist-chartist switching, α_1 and α_2 are sensitivity measures with respect to the majority and trend components.

The other switching probabilities, i.e. between the two chartist categories and fundamentalists, are defined as follows:[24]

$$\begin{aligned} \pi^{F+} &= \frac{v_2}{n}\left(\frac{N_C^+}{N}\,e^{U_{2,1}}\right);\ \pi^{+F} = \frac{v_2}{n}\left(\frac{N^F}{N}\,e^{-U_{2,1}}\right) \\ \pi^{F-} &= \frac{v_2}{n}\left(\frac{N_C^-}{N}\,e^{U_{2,2}}\right);\ \pi^{-F} = \frac{v_2}{n}\left(\frac{N^F}{N}\,e^{-U_{2,2}}\right), \end{aligned} \tag{14}$$

where model parameter $v_2 > 0$ is the frequency of this type of transition, N_C^+/N is the probability of meeting an optimistic chartist, etc.[25] The corresponding fitness measures $U_{2,1}$ and $U_{2,2}$ are in this case driven by the relative excess profits and are defined as:

$$\begin{aligned} U_{2,1} &= \alpha_3\left(\frac{r+\Delta P/v_2}{p} - R - s\,\frac{|P-F|}{P}\right) \\ U_{2,2} &= \alpha_3\left(R - \frac{r+\Delta P/v_2}{P} - s\,\frac{|P-F|}{P}\right), \end{aligned} \tag{15}$$

where parameter α_3 is the sensitivity of traders to differences in profits. The realised excess profits for optimistic chartists are given by $(r+\Delta P)/P - R$, where r are constant nominal dividends of the asset and R is the average real return from other investments ($R = r/F$), while the excess profit for pessimistic chartists is negative. On the other side, the expected excess profit of fundamentalists is given by the deviation from the fundamentals $s\,|P-F|/P$, where the parameter $0 < s < 1$ may be interpreted as a discount factor.

Finally, the new market price is set by means of a stochastic market impact function of total excess demand $ED = N_C\,D_C + N_F\,D_F$, with the speed of adjustment β and a small noise $\mu \sim \mathcal{N}(0, \sigma_\mu^2)$. The price change per time unit is restricted to a fixed tick size $\Delta P = \pm 0.01$, as described in Lux and Marchesi (2000), and takes place probabilistically in the direction of the excess demand according to:

$$\begin{aligned} \pi^{\uparrow P} &= \min(\max(0, \beta\,(ED+\mu)), 1) \\ \pi^{\downarrow P} &= \min(-\min(0, \beta\,(ED+\mu)), 1). \end{aligned} \tag{16}$$

[24]For example, the flow of fundamentalists to the optimistic chartists group is $N_F\,r^{F+} - N_C^+\,r^{+F}$, where $r^{F+} \sim \mathcal{N}(\pi^{F+}, 1/N_F)$ and $r^{+F} \sim \mathcal{N}(\pi^{+F}, 1/N_C^+)$.

[25]The switching probabilities are based on the assumption of direct interaction between agents (herding effect) and, thus, their values depend on the sizes of agent groups pursuing a common strategy.

Alternatively, the Lux model could be implemented at the individual (agent) level. In this approach, the evolution of the cluster sizes will have a different dynamics because the effective net transitions will be more affected by the tails of the sampling distributions than in the cluster-composite approach presented previously.

The Farmer-Joshi model

The Farmer-Joshi model is also built on two types of agent expectations, i.e. fundamentalism and chartism (see equation (17)), where $p_t = \log(P_t)$ and θ^i is a time lag uniformly distributed between θ_{min} and θ_{max}. The logarithmic fundamental value f_t follows a random walk $f_{t+1} = f_t + \eta_{t+1}$ with exogenous fundamental shocks $\eta_t \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2)$. Fundamentalist agents perceive a noisy fundamental value $f_t^i = f_t + \tilde{f}^i$ with an individual constant random offset $\tilde{f}^i$, assigned uniformly between $\tilde{f}_{min}$ and $\tilde{f}_{max}$, where $\tilde{f}_{min} = -\tilde{f}_{max}$ so that the entire range is $2\,\tilde{f}_{max}$.

$$E_{F,t}^i[\Delta p_{t+1}] = f_{t+1}^i - p_t;\ E_{C,t}^i[\Delta p_{t+1}] = p_t - p_{t-\theta^i} \tag{17}$$

In this model, the agents' strategies are state-dependent and non-linear by taking into consideration the current stock inventory and by adding individually-paired entry $T^i > 0$ and exit $\tau^i \leq 0$ thresholds, which are randomly initialized from uniform distributions ranging from T_{min} to T_{max} and from τ_{min} to τ_{max}, respectively.[26] An agent's asset placement size is given by $D^i = a\,(T^i - \tau^i)$, where the constant $a > 0$ is called scaled parameter for capital assignment. For example, if the mispricing in (17) is smaller than $-T$ and the agent carries no position yet, a fundamentalist enters a short position $-D$ which is hold until the mispricing goes above $-\tau$.

The running frequency of the model is daily, with no intraday microintervals. Another model feature is that no evolution is implemented, i.e. agents do not change strategies and thus the number of agents corresponding to each type remains constant $N/2$. Finally, the price discovery mechanism assumes the presence of a risk neutral market maker and takes the form of the price impact function of net demand in equation (18), where λ represents a liquidity scale factor and $\xi_t \sim \mathcal{N}(0, \sigma_\xi^2)$ is a random term corresponding to random perturbations in the price.

$$p_{t+1} = p_t + 1/\lambda \sum_{i=1}^{N} D_t^i + \xi_{t+1} \tag{18}$$

Table 1 presents an overview of the three previous ABMs with focus on the classification of their building blocks, as introduced in Section 2.

[26]Other conditions regarding the thresholds are $-T < \tau < T$ and $|\tau_{min}| \leq T_{min}$.

	Kirman	Lux	Farmer-Joshi
Model design	2-type	3-type	2-type
Agent design			
Expectation	behavioral & majority assessment	behavioral	behavioral
Fundam. process	noisy constant	constant	random walk
Risk preference	neutral	neutral	neutral
Wealth	n.a.	n.a.	asset position
Learning	probabilistic switching	evolutionary switching	n.a.
Price discovery	Walrasian tâtonnement	prob. discrete market impact	market impact with noise

Table 1: Comparative model classification

4.2 Calibration and output analysis

Since the dynamics of an ABM also depends on the selected set of parameters, in the following we will compare the outcomes of both the original parametrisations and a set of alternative parametrisations (marked with $\star$).[27] In the case of the Farmer-Joshi model we are using and ad-hoc parametrisation instead of the optimised one as the alternative choice, because the calibration procedure identifies a configuration where no trading occurs and the price is driven only by the noisy fundamental value (see Table 4). The calibration procedure is based on the simulated indirect inference method introduced in Gilli and Winker (2003), which minimizes the aggregated distance between a selected set of real data moments m^e and the simulated moments generated by the actual ABM, given its current parametrsiation, denoted by $m^s|\theta$. The distance measure is operationalized through the objective function introduced in Winker et al. (2007), where the weighting matrix $\hat{W}(m^e)$ for all pairs of moments is derived by inverting the variance-covariance matrix of the joint distribution of moments obtained from a block-bootstrap analysis. The objective function $f(\theta)$ takes the following form:

$$f(\theta) = \left[\frac{1}{I}\sum_{i=1}^{I}[(m_i^s|\theta) - m^e]\right]' \hat{W}(m^e) \left[\frac{1}{I}\sum_{i=1}^{I}[(m_i^s|\theta) - m^e]\right] \tag{19}$$

In this paper, we are using the same empirical exchange rate time-series (DEM/ USD, 1991–2000) and include the same moments as in Winker et al. (2007) and Jeleskovic (2011): the mean and standard deviation of log-returns, the L2-distance in the Lilifors statistic, the mean over the 5-10% thresholds for the Hill estimate of the right tail index, the sum of ARCH(10) coefficients, the Geweke & Porter-Hudak (GPH) estimator of the degree of fractional integration, and the OLS regression coefficients of the Augmented Dickey-Fuller

[27]There is an exception in the case of the Kirman model where, because of missing original parameters, we use the parameters presented in Winker et al. (2007)

test with drift and no lags (for more details see also Winker and Jeleskovic (2006, 2007)). The benchmark daily DEM/ USD time-series counts 2285 observations and is plotted in Fig. 1 and the estimate of the weighting matrix $\hat{W}$ is given by:

$$\hat{W} = \begin{pmatrix} 1.0 \cdot 10^8 & 1.4 \cdot 10^6 & 332.19 & 3.2 \cdot 10^3 & 4.9 \cdot 10^4 & -4.2 \cdot 10^4 & -3.5 \cdot 10^6 \\ 1.4 \cdot 10^6 & 8.7 \cdot 10^6 & -193.73 & 4.6 \cdot 10^3 & 788.00 & 4.8 \cdot 10^3 & 1.4 \cdot 10^5 \\ 332.19 & -193.73 & 1.5228 & -0.4122 & 0.9262 & -0.4249 & -42.717 \\ 3.2 \cdot 10^3 & 4.6 \cdot 10^3 & -0.4122 & 15.705 & 28.963 & 11.942 & -255.98 \\ 4.9 \cdot 10^4 & 788.00 & 0.9262 & 28.963 & 245.10 & -64.102 & -442.80 \\ -4.2 \cdot 10^4 & 4.8 \cdot 10^3 & -0.4249 & 11.942 & -64.102 & 217.29 & -126.67 \\ -3.5 \cdot 10^6 & 1.4 \cdot 10^5 & -42.717 & -255.98 & -442.80 & -126.67 & 5.2 \cdot 10^5 \end{pmatrix}$$

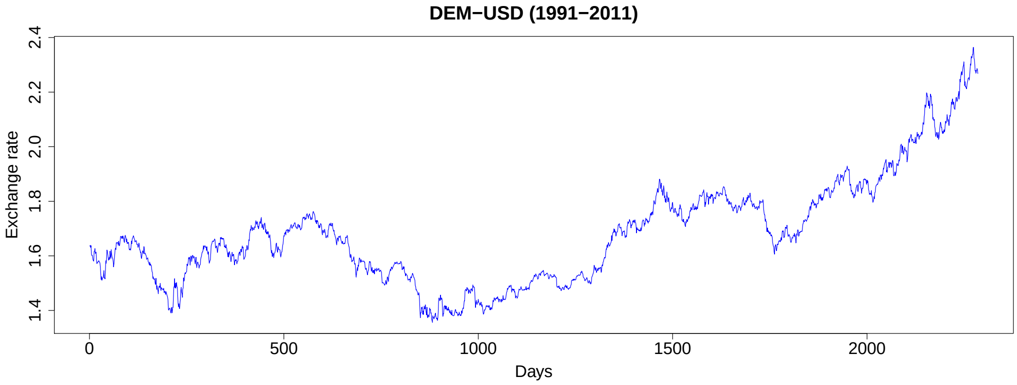

Figure 1: DEM/USD time series (1991–2000)

Finally, the optimisation problem is approached using a standard Genetic Algorithm with a population of 200 individual solutions, 30 generations, a crossing probability of 25% and a mutation probability per bit of 5%. The simulated moments of each individual solution are computed as the mean over 100 different price time-series generated by the particular solution, starting with different random seeds. Moreover, the entire Genetic Algorithm is restarted ten times with different random seeds. Afterwards, the overall best solution is chosen. The imposed search intervals for the optimisation process are presented in Appendix B. Tables 2, 3 and 4 contain the values of the parameters for both the (unoptimised) original and the optimised cases, as well as the associated mean values of the objective function, of the approximate entropy of log returns $ApEn(1, 0.2\,\sigma)$ and of the Higuchi fractal dimension of the price time-series, computed over 100 runs with various random seeds.[28] For comparison, the approximate entropy of the original DEM/USD data is 2.2245, while the normalized approximate entropy is 203.82 and the Higuchi fractal dimension is 1.4645.

[28]We are using the set-up provided in Pincus and Kalman (2004) where $m = 1$ and r equals 20% of the time-series standard deviation.

Moreover, as benchmark for the ABM models, we have calibrated a random walk model, with drift, of the following form:

$$P_{t+1} = P_t + \mu + \xi, \tag{20}$$

where $\xi \sim \mathcal{N}(0, \sigma^2)$, $P_0 = 1.681$, $P_t \geq 0.001$ and the optimised parameters are $\mu = 0.0007$, $\sigma = 0.0133$, with an average fit $f(\theta) = 73.592$. The mean value over 100 different runs for the approximate entropy of the calibrated random walk is 2.2837, for the normalized approximate entropy 171.72 and for the Higuchi fractal dimension 1.4946. The absolute *ApEn* shows that the random walk series is more irregular than the original data, while the normalized *ApEn*, on the contrary, indicates more regularity. On the other side, the values of the Higuchi fractal dimension index (*HFD*) are quite close to each other, in an interval between 1.45 and 1.50.

Regarding the model fit, an important observation is that only the optimised Lux model is better than the calibrated random walk, while the Farmer-Joshi and, especially, the Kirman models are worse. Also, even if not part of the set of moments targeted by the objective function $f(\theta)$, the calibrated Lux model is the closest to the original time-series regarding the absolute *ApEn* and *HFD*, while the normalized *ApEn* is even larger than the one describing the random walk process. The Farmer-Joshi model provides the second best fit, which is also in line with the ranking given by the absolute *ApEn*. However, it would be ranked first based on the normalized *ApEn* and *HFD*. The Kirman models are invariantly the worst performing according to all measures, but to different extent.

Label	Interpretation	Parameter values			
		$K_{\texttt{ag}}$		$K_{\texttt{cl}}$	
		original	optim. ($\star$)	original	optim. ($\star$)
α_F	Fundamentalist adjustement speed	0.08	0.04	0.08	0.02
c	Market maker price adjustement speed	0.50	0.50	0.50	0.50
σ_q	Majority assessement noise	0.07	0.01	0.07	0.01
δ	Conviction probability	0.35	0.75	0.35	0.50
ε	Mutation probability	0.01	0.02	0.01	0.02
Objective fit $f(\theta)$		149,088	115,133	147,451	116,128
Approximate entropy of log returns		1.8257	1.7439	1.8354	1.7499
Normalized approximate entropy of log returns		3730	2121	3779	2168
Higuchi fractal dimension of price		1.6994	1.5419	1.6975	1.5450

Table 2: Parameters of the individual-based ($K_{\texttt{ag}}$) and cluster-based ($K_{\texttt{cl}}$) Kirman models (agents number $N = 100$)

Fig. 2 plots the price and fundamental value dynamics of a single realisation, generated by the previous ABM designs, both for the original (un-optimised) and the alternative parametrisations. In contrast to the original time series shown in Fig. 1, the length of each time series in Fig. 2 is only

Label	Interpretation	Parameter values	
		original	optimised (⋆)
ΔP	Fixed value of price increase change	0.01	0.002
α_F	Fundamentalist adjustement speed	0.01	0.01
t^C	Fixed amount assets bought or sold by a chartist	0.02	0.009
ν_1	Frequency of the in-between chartist switching	3.00	2.10
α_1	Chartist sensitivity measures w.r.t. majority opinion	0.60	1.60
α_2	Chartist sensitivity measures w.r.t. trend	0.20	0.40
ν_2	Frequency of the fundamentalist-chartist switching	2.00	1.60
α_3	Sensitivity of traders to differences in profit	0.50	0.10
s	Fundamentalist discount factor	0.75	0.75
R	Average real return from other investments	0.0004	0.0003
β	Speed of price adjustment	6.00	4.60
σ_μ	Noise std. dev.	0.05	0.065
Objective fit $f(\theta)$		65,149	26.64
Approximate entropy of log returns		2.4166	2.2380
Normalized approximate entropy of log returns		495	378
Higuchi fractal dimension of price		1.5323	1.3980

Table 3: Parameters of the Lux model (agents number $N = 500$)

Label	Interpretation	Parameter values		
		original	optimised	ad-hoc (⋆)
$T_{\min}$	Min. threshold for entering positions	0.2	0.46	0.1
$T_{\max}$	Max. threshold for entering positions	4.0	0.78	1.5
$\tau_{\min}$	Min. threshold for exiting positions	-0.2	-0.8	-0.1
$\tau_{\max}$	Max. threshold for exiting positions	0.0	0.0	0.0
a	Scale parameter for capital assignment	0.0025	0.0002	0.0025
$\tilde{f}_{\min}$	Min. offset for log of perceived fund. value	-2.0	-1.2	-0.1
$\tilde{f}_{\max}$	Max. offset for log of perceived fund. value	2.0	1.6	0.1
$\theta_{\min}$	Min. time delay for trend followers	1	1	1
$\theta_{\max}$	Max. time delay for trend followers	100	20	100
μ_η	Mean of fundamental shocks	0.0	-0.6	0.0
σ_η	Std. dev. of fundamental shocks	0.010	0.012	0.005
λ	Liquidity	1.0	1.5	0.5
σ_ξ	Noise driving price formation process	0.35	0.01	0.01
Objective fit $f(\theta)$		$122 \cdot 10^4$	2,411	2,541
Approximate entropy of log returns		2.2793	2.2822	2.2809
Normalized approximate entropy of log returns		6.41	228	225
Higuchi fractal dimension of price		1.5086	1.5011	1.5142

Table 4: Parameters of the Farmer-Joshi model (population size $N = 1200$)

three years (252 time periods per year), which follows a burn-in period of one year. The plots related to Kirman's model present a high degree of similarity, showing and alternation between periods when the price closely follows the fundamental values and periods of larger volatility. As opposed to the original parametrisation of the Lux model, the optimised Lux parameters generate a price series which does not exhibit a fast reversal to the fundamental price. Nevertheless, under this calibrated parametrisation, the Lux* model provides the best fit to the real DEM/USD time-series. The original parametrisation of the Farmer-Joshi model is associated with the worst fit of the objective function, its price outcome does not follow the fundamental value and also exhibits huge volatility clusters. On the other side, the alternative ad-hoc parametrisation generates a more "temperate" volatility, but the price still does not revert to the underlying fundamentals.

4.3 Complexity assessment

In this subsection, the complexity measures introduced in Section 2 are provided for all three models described in Subsection 4.1, both with the original parameters and under the alternative parametrisation previously discussed in Subsection 4.2 (marked with $\star$). Table 5 includes the average results over 100 runs with various random seeds, as well as the mean values of the objective function $f(\theta)$ defined in Winker et al. (2007) – since the optimisation is a minimsation problem, the lower the objective function value, the better the fit. Also, for the first two classes of complexity, i.e. difficulty of description and difficulty of creation, a lower value of the indicator is associated with a lower complexity. The less complex model–parametrisation pair is marked in Table 5 with boldface for each of the measures of complexity considered.

All implementations are built up on the same framework in order to enforce a common design structure, which closely follows the building blocks in Section 2. Thus, the framework defines abstract classes and their dependencies, e.g., *Agent/Cluster*, *Expectation*, *RiskPreference*, *MarketClearing*, *Evolution*, *WorldState* and *AgentBasedModel*, which have to be extended by each model implementation. The source files of the framework sum up to 159 lines of code and are added to the model specific code when computing $\mathcal{C}_{code}$. For computing $\mathcal{C}_{space}$, we use the class *Runtime methods* which relates to the memory consumption in the Java virtual machine.[29]

The Lux model has the lengthiest implementation mainly because of its lengthy description of the evolution process, associated also with a worse time and space performance – five times slower than the cluster-based Kirman while the number of iterations is only double due to the higher number of intraday microintervals. Also, the maximum cyclomatic complexity number records its highest value in the case of the Lux model. On the other side,

[29]Every Java application has a single instance of class java.lang.Runtime that allows the application to interface with the environment in which the application is running, including the virtual memory space assigned by the operating system to the Java process.

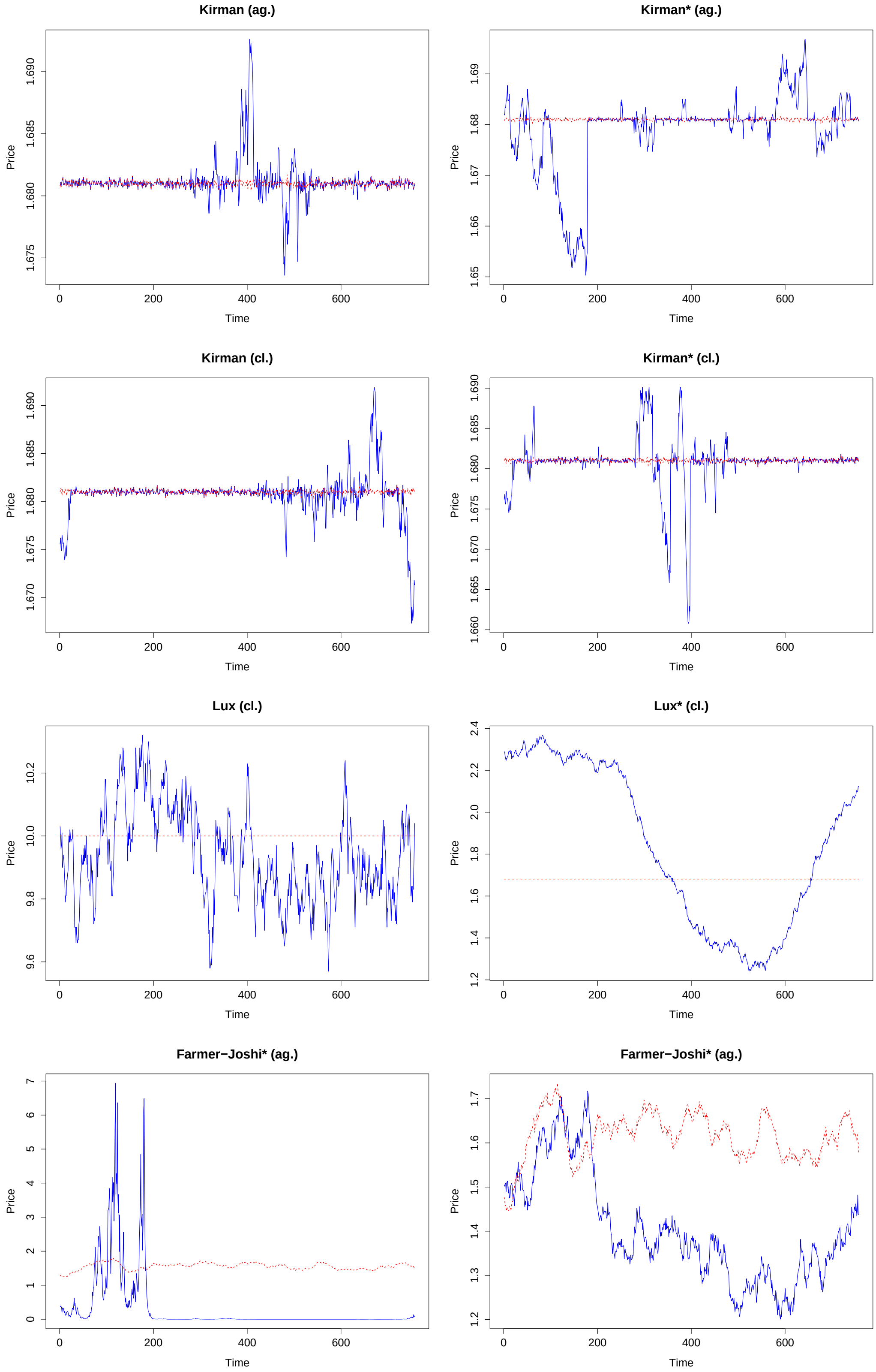

Figure 2: Fundamental value (dashed-red) and price (blue) time series

	$K_{\mathtt{ag}}$	$K^{\star}_{\mathtt{ag}}$	$K_{\mathtt{cl}}$	$K^{\star}_{\mathtt{cl}}$	Lux	Lux*	F-J	F-J*
ABM type class	ag.	ag.	cl.	cl.	cl.	cl.	ag.	ag.
Difficulty of description								
$\mathcal{C}_{\#P}$	**7**	**7**	**7**	**7**	13	13	14	14
$\mathcal{C}_{code}$ (lines)	373	373	399	399	579	579	**349**	**349**
$\mathcal{C}_{cyc-max}$	13	13	**12**	**12**	18	18	15	15
$\mathcal{C}_{cyc-av}$	3.07	3.07	**2.28**	**2.28**	3.00	3.00	3.08	3.08
$\mathcal{C}_{depth-max}$	6	6	**5**	**5**	**5**	**5**	6	6
$\mathcal{C}_{depth-av}$	2.10	2.10	**1.95**	**1.95**	2.05	2.05	2.06	2.06
Difficulty of creation								
$\mathcal{C}_{time}$ (ms)	1242	1217	57	38	250	264	30	**24**
$\mathcal{C}_{space}$ (MB)	1.94	1.95	1.95	1.96	3.74	3.74	**0.44**	**0.44**
Degree of organisation								
$\mathcal{C}_{ApEn(1,0.2\,\sigma)}$	0.53	0.37	0.50	0.40	0.92	**0.99**	1.00	0.44
$\mathcal{C}_{ApEn(1,0.2\,\sigma)/\sigma}$	1.31	0.77	1.24	0.81	5.71	**6.80**	4.64	1.60
$\mathcal{C}_{ApEn(2,0.2\,\sigma)}$	0.50	0.33	0.48	0.36	0.91	**0.97**	0.76	0.41
$\mathcal{C}_{ApEn(2,0.2\,\sigma)/\sigma}$	1.22	0.69	1.18	0.74	5.64	**6.67**	3.47	1.51
$\mathcal{C}_{HFD}$	1.551	1.511	**1.504**	**1.505**	1.550	1.550	1.897	1.950
Output								
Objective fit (E+03)	149	115	147	116	65	0.027	1220	2.51

Table 5: Complexity measures (100 runs with $T = 2350$ days each) [a]

[a] The number of iterations differs because of the various number of intraday microintervals, e.g., the Kirman (K) model has $2350 * 50$ time-steps, the Lux model $2350 * 100$ time-steps and the Farmer-Joshi (F-J) model only 2350 time-steps.

because of its missing evolution component and lacking intraday frequency, the Farmer-Joshi model is the fastest and also has the lowest memory consumption. As expected, the cluster-based Kirman is much faster than its individual-based version, since the evolution process is more compact, however the necessary memory space remains the same. Another observation regarding the first two complexity classes is that $\mathcal{C}_{time}$ is the only measure which is sensitive to a change in model parameters. Except for the Lux model, all other models record a decrease of the running time, ranging from 5% to 33%, for the calibrated configurations as compared to the standard parametrization.

Before discussing the third complexity class, i.e. degree of organisation, we describe how the internal information time-series mentioned in Subsection 3.3, which are the basis for computing the approximate entropy measure $\mathcal{C}_{ApEn}$, are obtained. For the cluster-based Kirman and Lux models, $\mathcal{C}_{ApEn}$ can be applied straight forward to the cluster weight dynamics of fundamentalists (since the weight of chartists is only complementary). For the atomic implementation of the Kirman model, a sequence corresponding to the percentage of "acting" fundamentalists can also be easily generated. In the case of Farmer-Joshi the number of agents in each category remains constant, however $\mathcal{C}_{ApEn}$ can be computed over the relative trading volume of fundamentalists given by $\sum_{i\in\mathcal{F}} |D_t^i| / \sum_{i=1}^{N} |D_t^i|$. These middle-layer time-series used as a proxy for self-organisation are depicted, for a single run, in Fig. 3. It can be observed that these time-series describe different dynamics depending both on the model and the calibration. For the Kirman models, the switching between the extreme cases, when all traders are of one single type, occurs more often when the model parameters are calibrated with the Genetic Algorithm as compared to the original parametrisation. In the case of the Lux model, there is a shift in the mean of the percentage of fundamentalists – the average fundamentalists weight is 0.77 with the original and 0.55 with the optimised parameters. A particular situation is observable in the case of the Farmer-Joshi model with original parameters, when after 230 days the fundamentalists do not participate in trading anymore.

As previously described in Subsection 3.3, the approximate entropy is lower for regular systems, larger for irregular and somewhere in-between for complex ones. At one extreme, a deterministic alternating sequence corresponds to a value of zero for $ApEn(1, 0.2\,\sigma)$. At the other end, a purely stochastic series of random draws from a normal distribution has an $ApEn$ of 2.30 and a normalized $ApEn$ of 2.31. Alternatively, a binomial distribution leads to an $ApEn$ of 0.70 and a normalized $ApEn$ of 1.39. In the following, we will quantitatively interpret the complexity associated with $\mathcal{C}_{ApEn(1,0.2\,\sigma)}$ by its proximity to a mid-value depending on the distribution type. The middle-layer distributions introduced in Subsection 3.3 resemble a binary distribution for the Kirman and Farmer-Joshi models (observable also from Fig. 3), and are thus benchmarked to a middle valued $ApEn$ of 0.35 and a middle normalized $ApEn$ of 0.70. In the case of the Lux model, the internal

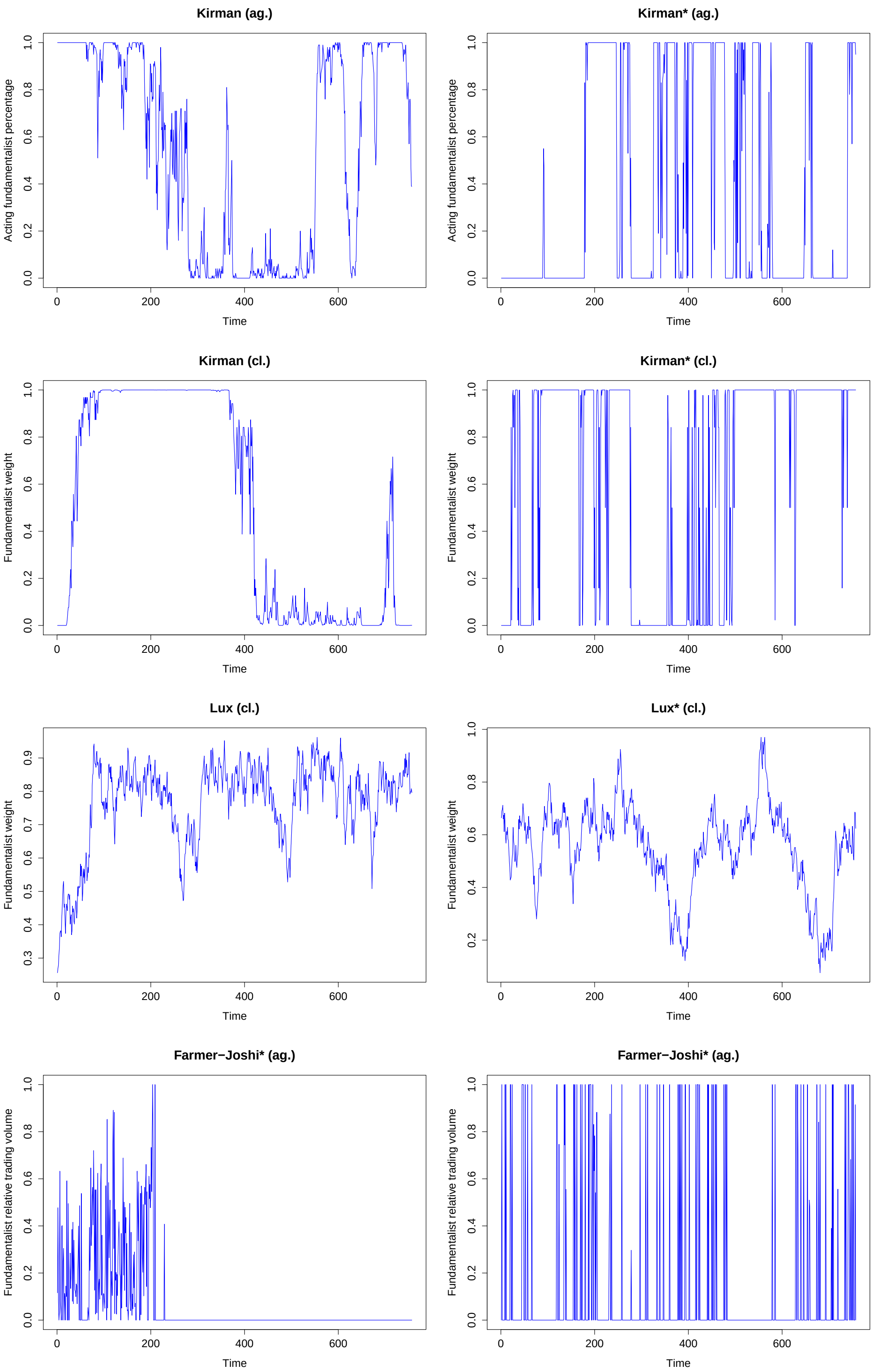

Figure 3: Internal information time series

distribution is Gaussian and will be benchmarked to a middle $ApEn$ and a middle normalized $ApEn$ of 1.15. Thus, two types of less complex systems can be identified, either too regular or too random. The entries marked in boldface in Table 5 correspond to the model exhibiting the largest absolute deviation from the middle value. For this comparison, only the calibrated models are considered.

The standard $ApEn(1, 0.2\,\sigma)$ measure shows that all considered ABMs exhibit a deterministic chaotic dynamics of the endogenous fraction of fundamentalists, far away from pure randomness and from periodic behavior. On the contrary, the normalized $ApEn$ identifies the Lux* and the Farmer-Joshi* as highly irregular. According to both the standard and the normalized $ApEn$, the less complex is the Lux* model – however the standard $ApEn$ captures a deviation from the benchmark mid-value towards more regular, while the normalized $ApEn$ towards more irregular. For the same parametrisation, i.e. original parameters, the cluster-based Kirman is slightly more regular than its individual-based alternative, which is explainable by the loss of heterogeneity due to aggregation. The values of $\mathcal{C}_{ApEn}$ are sensitive to the model parameters, meaning that the parameter choice influences the dynamics of the fundamentalist-chartist fractions. However, the direction of change is not consistent with the values of the objective function, e.g., an improvement of fit is associated with a $\mathcal{C}_{ApEn}$ increase for Lux and with a decrease in the case of Kirman and Farmer-Joshi models. On the other side, it is not necessary to have a direct relation between $\mathcal{C}_{ApEn}$ and the values of the objective function, since the internal information of the model is processed into the final output only by passing through the ABM specific design. In other words, it might be that for some models a more chaotic dynamics of agent type shares might correspond to a better fit, while for other models it might be exactly the other way around. Also, as previously noted, these middle-layer distributions can be different for various ABMs, i.e. binomial for the Kirman and Farmer-Joshi models and Gaussian for the Lux model.

Regarding the complexity captured by the Higuchi fractal dimension of the dynamics of fundamentalists fraction, i.e. $\mathcal{C}_{HFD}$, it can be noticed that the Farmer-Joshi model is the most complex, while the other three have similar values; by a small difference, the Kirman cluster could be ranked as the least complex model. Also, in the case of the Farmer-Joshi model, the calibration is associated with an increase in complexity, while for the atomic implementation of the Kirman model, it is the other way around. The Lux model and the cluster-based Kirman are mostly insensitive in this respect. However, the same remark as with the $ApEn$ interpretation should be taken into account – there is a tight coupling between the internal information of the model and the model design which, on one side could allow the identification of some complexity footprint of the model, but on the other side shortcuts the relation to the final output and, consequentially, to the model fit measure.

5 Conclusion and Outlook

We have identified and described the main building blocks, which are characteristic to most agent-based models for financial markets, as following: agent design, agent evolution and price discovery. These building blocks have been conceptually specified and exemplified with the main historical milestones and various derivations. Moreover, the previously introduced building blocks are connected within a framework which can be used as a common standard for developing and characterizing ABMs. The procedure is exemplified with three popular ABMs for financial markets: the Kirman model introduced in Kirman (1991, 1993), the Lux model described in Lux (1995, 1998) and Lux and Marchesi (1999, 2000), and the Farmer-Joshi model from Farmer and Joshi (2002). In the case of the Kirman model, two possible implementations are discussed: an individual-based (atomic) and a cluster-based design. In a next step, all three models are calibrated to a set of stylized facts computed on real data using the simulated indirect inference method introduced in Gilli and Winker (2003) and the objective function defined in Winker et al. (2007).

Alongside, we have started a discussion on how to define and measure size and complexity of ABMs. Similar to Lloyd (2001), we have considered three different dimensions of complexity and tried to identify and accommodate different measures of complexity which might be meaningful from an ABM perspective. Thus, the number of parameters, code length, cyclomatic complexity number and maximum block depth reflect the difficulty of description of an ABM, while the computational time and memory requirements deal with the difficulty of creation. A third class of complexity, i.e. the degree of organization, is meant to capture a different facet of complexity which is related to the dynamics of an ABM due to non-linear interactions, self-organisation and latent features of the system. We argue that the aggregate complexity of an ABM can be observed as a footprint by analyzing a middle-layer of information represented by the distributions of agent types. For 2-agent types models, e.g., the ones built on the fundamentalist-chartist dichotomy, we have assessed the complexity of this inner information time-series by means of approximate entropy and Higuchi fractal dimension.

The previously introduced complexity measures have been computed for the three selected ABMs, both under the original and under the calibrated parametrisations. None of the three models was uniquely identified as the most or the least complex by the entire range of proposed complexity measures, which shows that the various measures reflect different dimensions of complexity. For example, the atomic Kirman model was ranked as the least complex by one single measure out of thirteen, while the cluster-based Kirman models by six, the Lux model by five and the Farmer-Joshi by three. In fact, some measures of complexity, such as the ones included in the degree of organisation class, do not exhibit a strong relationship with the other measures which belong to the more static descriptive classes. Both aspects of complexity probably have to be considered jointly to gain proper insights in the overall "complexity" of a model in a broad sense. Further research is needed to gain a better understanding of the relationship between the two type of classes.

An extension of the current research could also analyse ABMs with autonomous-agents, where individual learning takes place by means of evolutionary algorithms, e.g., the Santa Fe Institute Artificial Stock Market. Also, it might be useful not to compare already settled, however too different ABM implementations, but instead to design new, incremental scenarios where small changes are applied to the individual building blocks, or single building blocks are replaced one at a time. For example, a possible experiment could target the effect of heterogeneity as follows. The base-model could be equipped with a pool of two alternative strategies (chartist and fundamentalist) with parameters which are fixed in time and common for all agents. The learning mechanism can consist in evolutionary stochastic switching, meaning that an agent will change his strategy based on the relative performance to an average benchmark. In the alternative case, the model could allow for a more diverse pool of strategies and also for developing new and improved ones using the Genetic Algorithm – the functional form of the forecasting rule should be common, e.g., a mix of chartist and fundamentalist components, but the parameters could be agent specific and the strategies pool could be frequently updated by means of the evolutionary algorithm. In order to maintain the same updating frequency as in the base-model, at every time step, the agents should make decisions about sticking to their current strategy or choosing a different one from the strategies pool, which is better fitted to the current market dynamics.

Another extension should deal with how to conduct model selection in a structured way, by taking into account both building blocks and complexity. After having provided some possible concepts and related measures of the complexity of an ABM, a link with its ability to reproduce stylized facts needs to be established. Intuition suggests that there should be a monotonic relation between the measure of complexity and the number of replicable stylized facts. However, a theoretical proof of this assumption might be hard to provide. Furthermore, even if the relationship turns out to be monotonic, the question arises how much complexity should be accepted in order to obtain a useful model and when improved replication turns into overfitting. In the context of (non)linear regression models, information criteria provide sensible answers. Obviously, the issue of matching stylized facts is strongly coupled with the problem of calibration or estimation of model parameters. For a given model design, it depends on the parameter values to what extent the model can replicate stylized facts or – in a more quantitative sense – fit moments of the real data. On the other side, as mentioned in Subsection 4.3, changing parameter values, e.g., in an optimization procedure, also directly influences the model's complexity, i.e. an ABM with the same design but different parameter settings exhibits different levels of complexity. Consequently, model selection has to take into account the complexity of the model for those parameter values obtained from calibration or estimation procedures.

A Higuchi Fractal dimension algorithm

Let $X(1)$, $X(2)$, ..., $X(N)$ be the finite set of N observations to be analyzed. A number of k new series X_k^m can be constructed as follows:

$$X_k^m = X(m), X(m+k), X(m+2\,k), \ldots, X\left(m + \left[\frac{N-m}{k}\right]\,k\right), \quad (21)$$

with $m = 1, 2, \ldots, k$ representing the initial values, k indicating the interval time (delay) and $[\cdot]$ denoting the Gauss notation, i.e. the integer part.

For each of the previous X_k^m series, the length of the curve can be computed as follows:

$$L_m(k) = \frac{1}{k}\left(\sum_{i=1}^{[(N-m/k)]} |X(m+i\,k) - X(m+(i-1)\,k)|\right)\frac{N-1}{\left[\frac{N-m}{k}\right]\,k}, \quad (22)$$

where the last factor is a normalisation factor.

The average length for each scale $k = 1, 2, \ldots, k_{max}$, $L(k)$, is computed as the mean of $L_m(k)$ where $m = 1, 2, \ldots, k$ and is proportional to k^{-D}, i.e. the average lengths follow a power law. D is considered the fractal dimension computed by Higuchi's algorithm and a least squares estimate for D is given by the slope of the line that fits the pairs $(\ln(L(k)), \ln(1/k))$.

B The search intervals for the calibration procedures

Label	Interpretation	Search intervals	Precision
α_F	Fundamentalist adjustement speed	[0.01, 0.64]	0.01
c	Market maker price adjustement speed	[0.5, 0.6]	0.1
σ_q	Std. dev. of the majority assessement noise	[0.01, 0.32]	0.01
δ	Conviction probability	[0.05, 0.80]	0.05
ε	Mutation probability	[0.01, 0.32]	0.01

Table B1: Search intervals and precision for the calibration of individual-based ($K_{\mathtt{ag}}$) and cluster-based ($K_{\mathtt{cl}}$) Kirman models

Label	Interpretation	Search interval	Precision
ΔP	Fixed value of price increase change	[0.001, 0.016]	0.001
α_F	Fundamentalist adjustement speed	[0.005, 0.020]	0.005
t^C	Fixed amount assets bought or sold by a chartist	[0.001, 0.032]	0.001
ν_1	Frequency of the in-between chartist switching	[0.1, 3.2]	0.1
α_1	Chartist sensitivity measures w.r.t. majority opinion	[0.1, 1.6]	0.1
α_2	Chartist sensitivity measures w.r.t. trend	[0.1, 0.8]	0.1
ν_2	Frequency of the fundamentalist-chartist switching	[0.1, 3.2]	0.1
α_3	Sensitivity of traders to differences in profit	[0.1, 1.6]	0.1
s	Fundamentalist discount factor	[0.65, 0.80]	0.05
R	Average real return from other investments	[0.0001, 0.0016]	0.0001
β	Speed of price adjustment	[2.1, 8.4]	0.1
σ_μ	Noise std. dev.	[0.005, 0.080]	0.005

Table B2: Search intervals and precision for the calibration of the Lux model

Label	Interpretation	Search interval	Precision
$T_{\min}$	Minimum threshold for entering positions	[0.01, 0.64]	0.01
$T_{\max}$	Maximum threshold for entering positions	[0.75, 3.90]	0.05
$\tau_{\min}$	Minimum threshold for exiting positions	[-0.8, -0.1]	0.1
$\tau_{\max}$	Maximum threshold for exiting positions	[0.0, 0.7]	0.1
a	Scale parameter for capital assignment	[0.0005, 0.0080]	0.0005
$\tilde{f}_{\min}$	Minimum offset for log of perceived fund. value	[-3.2, 0.1]	0.1
$\tilde{f}_{\max}$	Maximum offset for log of perceived fund. value	[0.1, 3.2]	0.1
$\theta_{\min}$	Minimum time delay for trend followers	[1, 8]	1
$\theta_{\max}$	Maximum time delay for trend followers	[50, 200]	10
μ_η	Mean of fundamental shocks	[-0.7, 0.8]	0.1
σ_η	Std. dev. of fundamental shocks	[0.001, 0.032]	0.001
λ	Liquidity	[0.1, 1.6]	0.1
σ_ξ	Noise driving price formation process	[0.01, 0.64]	0.02

Table B3: Search intervals and precision for the calibration of the Farmer-Joshi model

References

Akaike, H. (1974). A new look at the statistical model identification, *Automatic Control, IEEE Transactions on* **19**(6): 716–723.

Arthur, W. B. (2014). *Complexity and the Economy*, Oxford University Press.

Arthur, W. B., Holland, J., LeBaron, B., Palmer, R. and Tayler, P. (1997). Asset pricing under endogenous expectations in an artificial stock market. SFI Paper 96-12-093, Economic Notes, 1997. Reprinted in The Economy as an Evolving Complex System II. Edited (with S. Durlauf and D. Lane), Addison-Wesley, 1997.

Barde, S. (2015). A practical, universal, information criterion over nth order markov processes.

Brock, W. and Hommes, C. (1997). A rational route to randomness, *Econometrica* **65**(5): 1059–1095.

Brock, W. and Hommes, C. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model, *Journal of Economic Dynamics and Control* **22**(8-9): 1235–1274.

Chen, S.-H. (2012). Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective, *Journal of Economic Dynamics and Control* **36**(1): 1–25.

Chen, S.-H., Chang, C.-L. and Du, Y.-R. (2012). Agent-based economic models and econometrics, *Knowledge Engineering Review* **27**(2): 187–219.

Chen, S.-H. and Yeh, C.-H. (2001). Evolving traders and the business school with genetic programming: A new architecture of the agent-based artificial stock market, *Journal of Economic Dynamics and Control* **25**(3): 363–393.

Chen, S.-H. and Yeh, C.-H. (2002). On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis, *Journal of Economic Behavior & Organization* **49**(2): 217–239.

Chiarella, C., Dieci, R. and He, X. (2009). Heterogeneity, market mechanisms, and asset price dynamics, *in* T. Hens and K. Schenk-Hoppé (eds), *Handbook of Financial Markets: Dynamics and Evolution*, North-Holland, Amsterdam, pp. 277–344.

Chiarella, C. and Iori, G. (2002). A simulation analysis of the microstructure of double auction markets, *Quantitative Finance* **2**(5): 346–353.

Chiarella, C., Iori, G. and Perelló, J. (2009). The impact of heterogeneous trading rules on the limit order book and order flows, *Journal of Economic Dynamics and Control* **33**(3): 525–537.

Cliff, D. and Bruten, J. (1997). Minimal-intelligence agents for bargaining behaviors in market-based environments, *Technical report*, School of Cognitive and Computing Sciences. University of Sussex.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues, *Quantitative Finance* **1**: 223–236.

Cristelli, M., Pietronero, L. and Zaccaria, A. (2011). Critical overview of agent-based models for economics. arXiv preprint arXiv:1101.1847.

Crutchfield, J. P. and Young, K. (1989). Inferring statistical complexity, *Phys. Rev. Lett.* **63**: 105–108.

Daniel, G. (2006). *Asynchronous simulations of a limit order book*, PhD thesis, University of Manchester.

Du, D.-Z. and Ko, K.-I. (2014). *Theory of Computational Complexity*, John Wiley & Sons.

Fabretti, A. (2013). On the problem of calibrating an agent based model for financial markets, *Journal of Economic Interaction and Coordination* **8**(2): 277–293.

Farmer, J. D. and Joshi, S. (2002). The price dynamics of common trading strategies, *Journal of Economic Behavior & Organization* **49**(2): 149–171.

Fischer, T. and Riedler, J. (2014). Prices, debt and market structure in an agent-based model of the financial market, *Journal of Economic Dynamics and Control* **48**: 95 – 120.

Franke, R. (2009). Applying the method of simulated moments to estimate a small agent-based asset pricing model, *Journal of Empirical Finance* **16**(5): 804–815.

Frankel, J. A. and Froot, K. A. (1987). The dollar as an irrational speculative bubble: A tale of fundamentalisists and chartists. National Bureau of Economic Research Cambridge, Mass., USA.

Gilli, M. and Winker, P. (2003). A global optimization heuristic for estimating agent based models, *Computational Statistics & Data Analysis* **42**(3): 299–312.

Gode, D. and Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality, *Journal of Political Economy* **101**(1): 119–137.

Grazzini, J. and Richiardi, M. (2015). Estimation of ergodic agent-based models by simulated minimum distance, *Journal of Economic Dynamics and Control* **51**: 148–165.

Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression, *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 190–195.

Haughton, D. (1991). Consistency of a class of information criteria for model selection in non-linear regression, *Communications in Statistics - Theory and Methods* **20**: 1619–1629.

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory, *Physica D: Nonlinear Phenomena* **31**(2): 277–283.

Hommes, C. (2006). Heterogeneous agent models in economics and finance, *in* L. Tesfatsion and K. Judd (eds), *Handbook of computational economics*, Vol. 2, Elsevier/North-Holland, Amsterdam, pp. 1109–1186.

Hommes, C. and Wagener, F. (2009). Complex evolutionary systems in behavioral finance, *in* T. Hens and K. Schenk-Hoppé (eds), *Handbook of Financial Markets: Dynamics and Evolution*, North-Holland, Amsterdam, pp. 217–276.

Iori, G. and Porter, J. (2012). Agent-based modelling for financial markets, *Working Papers 12/08*, Department of Economics, City University, London.

Jeleskovic, V. (2011). *Agentbasierte Modelle für empirische Wechselkurse*, Peter Lang, Frankfurt am Main.

Kirman, A. (1991). Epidemics of opinion and speculative bubbles in financial markets, *in* M. Taylor (ed.), *Money and financial markets*, Blackwell, Cambridge, pp. 354–368.

Kirman, A. (1993). Ants, rationality, and recruitment, *The Quarterly Journal of Economics* **108**(1): 137–156.

LeBaron, B. (2000). Agent-based computational finance: Suggested readings and early research, *Journal of Economic Dynamics and Control* **24**(5): 679–702.

LeBaron, B. (2006). Agent-based computational finance, *Handbook of Computational Economics* **2**: 1187–1233.

LeBaron, B., Arthur, W. and Palmer, R. (1999). Time series properties of an artificial stock market, *Journal of Economic Dynamics and Control* **23**(9): 1487–1516.

LeBaron, B. and Winker, P. (2008). Introduction to special issue on agent based models in economic policy advice, *Journal of Economics and Statistics* **228**(2+3): 141–147.

Lloyd, S. (2001). Measures of complexity: a nonexhaustive list, *IEEE Control Systems Magazine* **21**(4): 7–8.

Lux, T. (1995). Herd behaviour, bubbles and crashes, *The Economic Journal* pp. 881–896.

Lux, T. (1998). The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions, *Journal of Economic Behavior & Organization* **33**(2): 143–165.

Lux, T. (2009). Stochastic behavioral asset pricing models and the stylized facts, *in* T. Hens and K. Schenk-Hoppé (eds), *Handbook of Financial Markets: Dynamics and Evolution*, North-Holland, Amsterdam, pp. 161–215.

Lux, T. and Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market, *Nature* **397**(6719): 498–500.

Lux, T. and Marchesi, M. (2000). Volatility clustering in financial markets: A microsimulation of interacting agents, *International Journal of Theoretical and Applied Finance* **3**(04): 675–702.

Mandelbrot, B. B. (1977). *The fractal geometry of nature*, WH Freeman and Co. New York.

Mandes, A. (Forthcoming). Microstructure-based order placement in a continuous double auction agent based model, *Algorithmic Finance* .

Manis, G. (2008). Fast computation of approximate entropy, *Computer Methods and Programs in Biomedicine* **91**(1): 48–54.

McCabe, T. J. (1976). A complexity measure, *IEEE Transactions on Software Engineering* (4): 308–320.

Mitchell, M. (2009). *Complexity: a guided tour*, Oxford University Press.

Nakamura, T., Judd, K., Mees, A. and Small, M. (2006). A comparative study of information criteria for model selection, *International Journal of Bifurcation and Chaos* **16**(8): 2153–2175.

Palmer, R., Arthur, W., Holland, J., LeBaron, B. and Tayler, P. (1994). Artificial economic life: a simple model of a stockmarket, *Physica D* **75**(1): 264–274.

Pincus, S. and Kalman, R. E. (2004). Irregularity, volatility, risk, and financial market time series, *Proceedings of the National Academy of Sciences of the United States of America* **101**(38): 13709–13714.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity., *Proceedings of the National Academy of Sciences* **88**(6): 2297–2301.

Pincus, S. M., Gladstone, I. M. and Ehrenkranz, R. A. (1991). A regularity statistic for medical data analysis, *Journal of Clinical Monitoring* **7**(4): 335–345.

Reuter, H., Breckling, B. and Jopp, F. (2011). Individual-based models, *in* F. Jopp, H. Reuter and B. Breckling (eds), *Modelling Complex Ecological Dynamics*, Springer, Heidelberg, pp. 163–178.

Rosenberg, L. H. (1998). Applying and interpreting object oriented metrics, Software Technology Conference (STC'98), Salt Lake City, Utah.

Schwarz, G. et al. (1978). Estimating the dimension of a model, *The Annals of Statistics* **6**(2): 461–464.

Tesfatsion, L. (2001). Introduction to the special issue on agent-based computational economics, *Journal of Economic Dynamics and Control* **25**(3): 281–293.

Watson, A. H., McCabe, T. J. and Wallace, D. R. (1996). Structured testing: A testing methodology using the cyclomatic complexity metric, *NIST special Publication* **500**(235): 1–114.

Westerhoff, F. (2009). Exchange rate dynamics: A nonlinear survey, *in* J. Rosser, J.B. (ed.), *Handbook of Research on Complexity.*, Edward Elgar, Cheltenham, UK, pp. 287–325.

Westerhoff, F. (2010). A simple agent-based financial market model: direct interactions and comparisons of trading profits, *Nonlinear Dynamics in Economics, Finance and Social Sciences* pp. 313–332.

Willems, F. M., Shtarkov, Y. M. and Tjalkens, T. J. (1995). The context-tree weighting method: basic properties, *Information Theory, IEEE Transactions on* **41**(3): 653–664.

Winker, P., Gilli, M. and Jeleskovic, V. (2007). An objective function for simulation based inference on exchange rate data, *Journal of Economic Interaction and Coordination* **2**: 125–145.

Winker, P. and Jeleskovic, V. (2006). The unconditional distribution of exchange rate returns: Statistics, robustness, time aggregation, *Technical Report WP008-06*, University of Essex, Colchester.

Winker, P. and Jeleskovic, V. (2007). Dependence of – and long memory in – exchange rate returns: Statistics, robustness, time aggregation, *Technical Report WP011-07*, University of Essex, Colchester.