

Lamperti, Francesco

**Working Paper**

## Empirical validation of simulated models through the GSL-div: An illustrative application

LEM Working Paper Series, No. 2016/18

**Provided in Cooperation with:**

Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies

*Suggested Citation:* Lamperti, Francesco (2016) : Empirical validation of simulated models through the GSL-div: An illustrative application, LEM Working Paper Series, No. 2016/18, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM), Pisa

This Version is available at:

<https://hdl.handle.net/10419/174527>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

INSTITUTE  
OF ECONOMICS



Scuola Superiore  
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics  
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy  
ph. +39 050 88.33.43  
institute.economics@sssup.it

# LEM

## WORKING PAPER SERIES

**Empirical Validation of Simulated Models  
through the GSL-div: an Illustrative Application**

Francesco Lamperti <sup>◦</sup>

<sup>◦</sup> Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy

**2016/18**

**April 2016**

**ISSN(ONLINE) 2284-0400**

# Empirical Validation of Simulated Models through the *GSL-div*: an Illustrative Application

Francesco Lamperti <sup>\*1</sup>

<sup>1</sup>*Institute of Economics and LEM, Scuola Superiore Sant'Anna*

April 14, 2016

## Abstract

A major concern about the use of simulation models regards their relationship with the empirical data. The identification of a suitable indicator quantifying the distance between the model and the data would help and guide model selection and output validation. This paper proposes the use of a new criterion, called *GSL-div* and developed in Lamperti (2015), to assess the degree of similarity between the dynamics observed in the data and those generated by the numerical simulation of models. As an illustrative application, this approach is used to distinguish between different versions of the well known asset pricing model with heterogeneous beliefs proposed in Brock and Hommes (1998). Once the discrimination ability of the *GSL-div* is proved, model's dynamics are directly compared with actual data coming from two major stock market indexes (EuroSTOXX 50 for Europe and CSI 300 for China). Results show that the model, once calibrated, is fairly able to track the evolution of both the two indexes, even though a better fit is reported for the Chinese stock market. However, I also find that many different combinations of traders behavioural rules are compatible with the same observed dynamics. Within this heterogeneity, an emerging common trait is found: to be empirically valid, the model has to account for a strong trend following component, which might either come from a unique trend type that heavily extrapolates information from past observations or the combinations of different types with milder, or even opposite, attitudes towards the trend.

**Keywords:** Simulated Models, Empirical Validation, Model Selection, *GSL-div*.

**JEL:** C15, C52, C63

---

<sup>\*</sup>*email:* [f.lamperti@sssup.it](mailto:f.lamperti@sssup.it) - The author acknowledges financial support from European Union's 7th FP for research, technological development and demonstration under G.A. No 603416 - Project IMPRESSIONS (Impacts and risks from high-end scenarios: Strategies for innovative solutions).

# 1 Introduction

Empirical validation is crucial for all modelling efforts that attempt at providing support to policy decisions, independently of their theoretical background. Even though Agent Based Models (ABMs) have often been advocated as promising alternatives to neoclassical models rooted in the dogmatic paradigms of rational expectations and representative agents, there are still some concerns about how to bring them *down* to the data (Windrum et al., 2007; Gallegati and Richiardi, 2009; Grazzini and Richiardi, 2015). In macroeconomics, for example, Giannone et al. (2006), Canova and Sala (2009) and Paccagnini (2009) provide details about how to estimate and validate Dynamic Stochastic General Equilibrium models. However, their approach cannot be extended to settings where an analytical solution of the model (or an equilibrium) does not exist, which are typical cases in ABMs, system dynamics and complex systems more in general. Broadly speaking, these numerical models are validated through a comparison of the statistical properties emerging from simulated and real data. In many cases, this amounts at replicating the largest possible number of stylized facts characterizing the phenomenon of interest (see Dosi et al., 2010, 2013, 2015 for business cycle properties, credit and interbank markets or Pellizzari and Forno, 2006; Jacob Leal et al., 2015 for financial markets). Recent attempts are trying to enrich empirical validation going beyond simple replication of empirical regularities, thereby requesting models to generate series that exhibit the same dynamics (Marks, 2013; Lamperti, 2015), conditional probabilistic structure (Barde, 2015) and causal relations (Guerini and Moneta, 2016) as those observed in the real world data. At least partially, such contributions have been motivated by the unsatisfactory results delivered by calibration. In general, it is difficult to justify the choice of one specification of model parameters over another, and calibration can be thought as the exercise of selecting the best values of the parameter set reproducing real data. In those cases where the model is sufficiently simple and well behaved, it is possible to derive a closed form solution for the distributional properties of a specific output of the model, and then to estimate the parameters governing such distributions accordingly (Alfarano et al., 2005, 2006; Boswijk et al., 2007). However, in the majority of cases, policy oriented models do not allow such procedures (see, for example, contributions in Dawid and Fagiolo, 2008 and LeBaron and Winker, 2008). When models' complexity prevents to obtain closed form solutions, more sophisticated techniques are required. Bianchi et al. (2007) and Bianchi et al. (2008) target a specific medium-scale macroeconomic agent based model and estimate some of its parameters by indirect inference (Gourieoux and Monfort, 1997). Starting from the same procedure, Gilli and Winker (2003) and Winker et al. (2007) introduced an algorithm and a set of statistics leading to the construction of an objective function used to estimate models of exchange rate and to push them closer to the properties of real data.<sup>1</sup> Recently, Grazzini and Richiardi (2015) proposed the approach of simulated minimum distance for estimation of ergodic ABMs, both in the long run equilibrium and during transitional dynamics, while Recchioni et al. (2015) used a simple gradient-based calibration procedure to conveniently sample the parameter space minimizing a standard loss function based on the cumulative squared errors. The key choice of a calibration exercise seems to boil down to the function that measures models' fit

---

<sup>1</sup>On the use of indirect inference see also Fabretti (2012).

with the data. According [Winker et al. \(2007\)](#), the moments and the statistics used in the objective function must be robust, reflect statistical properties of the real data and exhibit the potential to discriminate between alternative models or parameter values. Not all calibration procedures appears persuasive in this respect. For example, [Amilon \(2008\)](#) estimates a relatively simple model of financial markets with 15 parameters (but only 2 or 3 agents) by efficient method of moments and reports an high sensitivity of the model to the assumptions on the noise term and stochastic components, questioning the performance of calibration exercises more in general. In addition, it appears straightforward that even if calibration delivers one or more array of parameters that maximise model's fit with the data, it is not automatic that this fit is a reasonably good one. Therefore, one of the objectives of validation exercises should be the assessment of somehow calibrated models in reproducing, ex-post, various properties of the real data and, in addition, to inform about the behaviour of the model within large regions of the parameter space (see also [Grazzini et al., 2015](#)).<sup>2</sup>

In this paper I present an application of the *GSL-div* developed in [Lamperti \(2015\)](#) to validate model's output against real word data and explore the behaviour of the model quantifying the distance between the dynamics observed in the data and those numerically simulated. Validation is achieved capturing the ability of a given model to reproduce the distributions of time changes (that is, changes in the process' values from one point in time to another) in the real-world series, without the need to resort to any likelihood function or to impose requirements of stationarity. The *GSL-div* adds something that seems missing in the literature: a precise quantification of the distance between the model and data with respect to their dynamics in the time domain. On this side, my work builds on [Marks \(2013\)](#) and extend it by capturing and emphasizing the dynamical nature of time series models, which is, for example, loosely represented by the longitudinal moments used in many calibration exercises. The *GSL-div* is tested on the series produced by the well known asset pricing model with heterogeneous traders developed in [Brock and Hommes \(1998\)](#).

The rest of the paper is organized as follows. Section 2 introduces the *GSL-div*, discusses its main properties and provides a simple example; section 3 summarizes the mathematical structure of the model that will be used throughout the paper and validated against historical data; section 4 constitutes the core of this contribution, it illustrates and discusses the results I obtained. Finally, section 5 concludes the paper and provides some insights into future research.

## 2 Validation and the GSL-div

### 2.1 Validation

Validation is a complex task that encompasses diverse aspects of the overall modelling activity. In this paper I interpret validation as the exercise of assessing the fit of one or different models with empirical data. To provide a general context, [Manson \(2002\)](#) distinguishes between *output validation* and *structural validation*. The latter asks how well the simulation model represents the (prior) conceptual model of the real-world system, while the former asks how

---

<sup>2</sup>As will be briefly discussed below, I refer here to what is called *output validation*.

successfully the simulations' output exhibits the historical behaviours of the real-world target system. Output validation can be directly related to what [Leombruni et al. \(2006\)](#) define as *empirical validity* of a model, i.e. validity of the empirically occurring true value relative to its indicator. Following [Rosen \(1985\)](#), it is useful to think of two parallel unfolding: the evolution of the system (an economy, a market or whatever) and the evolution of the model of the system. If the model is correct, properly calibrated and initial conditions have been fixed according to the initial status of the real system, the simulation should mirror the historical evolution of such system with respect to the variables, or statistics, of interest.

In our context, a *model* is broadly defined as a representation of a system that is able to produce some synthetic output tracking the evolution of the system itself. Formally, the output of a model can be represented by the collection of all micro-states at time  $t$ ,  $\mathbf{X}_t \equiv \{x_{i,t}\}$  with  $i = 1, \dots, N$  and  $t = 1, \dots, T$  such that

$$x_{i,t} = f_i(\boldsymbol{\theta}, \mathbf{X}_{t-1}), \quad (1)$$

where  $f_i$  can be any (deterministic or stochastic) real valued function and  $\boldsymbol{\theta} \in \Theta \subset \mathbf{R}^d$  is a vector of  $d$  parameters. I assume that, for each model,  $\boldsymbol{\theta}$  and  $\mathbf{X}_0$  are exogenously given. In other words, the model has already been calibrated: initial conditions and parameters are assigned precise values. In addition, *real world data* are defined as the empirically observable elements of the system.

## 2.2 The *GSL-div*

The *GSL-div* is a measure, developed in [Lamperti \(2015\)](#), that determines the degree of similarity between the dynamics observed in real data and those produced by the numerical simulation of a model. The only input it requires are the real and simulated series. It should be noticed that such a comparison might involve objects having different dimensions. While real quantities can be observed once and only once, if the model is stochastic, many different realizations might be obtained varying the seed of the random number generator. To the contrary, if the model is deterministic, a unique (or many identical) series will be obtained for each variable of interest. The *GSL-div* can be used in both the two cases, but it has been thought to treat stochastic models, which I see as the most relevant cases, provided that many economic decision or events (e.g. innovation outcome) are inherently uncertain.

The estimation of the *GSL-div* follows a simple, three steps procedure that is proposed and briefly discussed here:

1. Time series (both real and simulated) are symbolized
2. Distributions of time-changes are estimated for windows of different lengths
3. The distance between distributions from real and simulated data are evaluated and aggregated.

The first step consists in series' symbolization. This procedure is carried out to constrain them to take only a finite set of values. Let  $\{x(t)\}_{t=1}^T$  be a time series of total length  $T$  where each  $x(t)$  is a real number. To symbolize it, I firstly take the real interval  $[x_{min}; x_{max}]$  and partition it in  $b \in \mathbf{N}_0$  subintervals, each of equal length. These intervals are numbered increasingly from 1

to  $b$ , with 1 assigned to  $[x_{min}; x_{min} + \frac{(x_{max}-x_{min})}{b})$ . The parameter  $b$  controls for the precision of the symbolization: for  $b = 1$  the symbolized series takes one and only one value (namely 1) while for  $b \rightarrow \infty$  we are back to the (scaled) real-valued process. The symbolization is simple and works as follows: each  $\{x(t)\}_{t=1}^T$  is mapped into the natural number corresponding to the partition interval where it falls. As an example, consider the following time series  $x(t)$  with  $T = 3$ :  $\{0; 0.4; 1\}$ . Choosing  $b = 2$ , the symbolized series will be  $x^s(t) = \{1, 1, 2\}$ , while choosing  $b = 9$  the symbolized series becomes  $x^s(t) = \{1, 4, 9\}$ , where the apex  $s$  stands for *symbolized*. Obviously, increasing  $b$  the information loss about the behaviour of the stochastic process due to the symbolization becomes smaller and smaller. However, I will show that the *GL-div* is very precise in recognizing similarity in time series dynamics even with very low values of  $b$ .

Now, I iteratively construct *blocks* of symbols having lengths  $l = 1, \dots, L$  with  $L \leq T$  by pasting together successive symbolized observations. Each of the resulting blocks with  $l > 1$  corresponds to a realized pattern of time to time changes in the process. For example, with  $b = 9$  and  $l = 1$  previous symbolized time series,  $x^s(t)$ , comprehends blocks  $\{1, 4, 9\}$ ; with  $l = 2$  it comprehends  $\{14, 49\}$ , indicating two increasing trends with the second being more pronounced than the first. Let  $S_{l,b} = \binom{A_b}{l}$  be an alphabet of symbols where  $A_b = \{1, 2, \dots, b\}$ . That is,  $S_{l,b}$  is the set of all the  $l$ -combinations of  $A_b$  and represents the set of all possible blocks with length  $l$ . The cardinality of the alphabet is defined as  $a_{l,b} = 2^{S_l} = b^l$ ,  $\forall l = 1, \dots, L$  and corresponds to the number of different symbols that blocks of length  $l$  might be associated to once  $b$  is chosen. It is relevant to see that, with  $l$  fixed, blocks might overlap. This is a relevant feature allowing to capture each possible pattern of time changes independently of the initial (and final) observation in the sample. If series are of length  $T$ ,  $T - l + 1$  blocks will be obtained for each value of  $l$ .  $L$  represents the maximum length of the windows which are used to compare the behaviour of the real data with the synthetic ones. It has to be chosen considering both (i) the nature of the phenomenon of interest and (ii) the size of the available real-world time series.

Now, frequency of symbols in each series are estimated. Let  $x^s(t)$  and  $y^s(t)$  be two symbolized time series.  $S_l$  is the alphabet at length  $l$  once  $b$  has been chosen, and  $\mathbf{f}, \mathbf{f}'$  are vectors collecting the occurrence frequencies of all available symbols. Similarity in the behaviour of the time series is inferred comparing symbols' frequency distributions obtained from the two time series for different *blocks's* length, which obviously corresponds to consider windows of different length to observe time to time changes. This allows, for example, to isolate and separate similarities at different time scales. For each value of  $l = 1, \dots, L$  the distance between  $\mathbf{f}_l$  and  $\mathbf{f}'_l$  is evaluated through the *L-divergence* (Lin, 1991). Their aggregation gives instead the *GSL-div* between couples of distributions.

**Definition 2.1.** (*GSL-div between distributions*)

Assume that  $b$  and  $L$  have been fixed. Let  $x^s(t)$  and  $y^s(t)$  be two symbolized time series exhibiting respectively frequencies  $\mathbf{f}_x$  and  $\mathbf{f}_y$  over the alphabet  $S_l$ . We define the Generalized Subtracted

*L*-divergence between the two distributions as

$$\begin{aligned}
D_{GSL}(\mathbf{f}_x || \mathbf{f}_y) &= \sum_{i=1}^L w_i \left( -2 \sum_{s \in S_i} m_i(s) \log_{a_i} m_i(s) + \sum_{s \in S_i} f_{y,i}(s) \log_{a_i} f_{y,i}(s) \right) \\
&= \sum_{i=1}^L w_i \left( 2H^{S_i}(\mathbf{m}_i) - H^{S_i}(\mathbf{f}_{y,i}) \right), \tag{2}
\end{aligned}$$

where the symbol  $H^{S_i}(\cdot)$  indicates the Shannon entropy of a distribution over the state space  $S_i$  and  $a_i$  the cardinality of the alphabet available at length  $l = i$ .

When dealing with a deterministic model, the use of (2) might be satisfactory in the characterization of the distance between real and simulated dynamics, since all the available information has already been exploited. In the more interesting case of stochastic models, in principle, one might want to estimate the distance between data and model relying on the probabilistic structure of the latter. For example, one would like to feed the *GSL-div* with the *true* probability that the model assigns to each sequence of symbols rather than its frequency. However, if the model is not solvable analytically, the only information about the stochastic process underlying the aggregate behaviour of these models is available through the synthetic series they produce. Now, imagine to take an ensemble of independent runs of the same, previously calibrated model. In this context, the following proposition is proved and discussed in details in Lamperti (2015, appendix B).

**Proposition 1.** *Let  $\bar{p}_\mu(s)$  be the average probability that model  $\mu$  assign to each symbol in the interval  $t \in [1, \dots, T]$  and  $p(s)$  the frequency of the same symbol observed in the real data. The *GSL-div* between  $\bar{p}_\mu(s)$  and  $p(s)$  is given by*

$$\begin{aligned}
GSL(p(s) || \bar{p}_\mu(s)) &= \sum_{i=1}^L w_i \mathbf{E} \left( -2 \sum_{s \in S_i} m_i(s) \log_{a_i} m_i(s) + \sum_{s \in S_i} f(s) \log_{a_i} f(s) \right) \\
&+ \sum_{i=1}^L w_i \left( \frac{B_i^m - C^m}{4T_i} - \frac{B_i^{\bar{p}_\mu} - C^{\bar{p}_\mu}}{2T_i} \right) + O(T^{-2}). \tag{3}
\end{aligned}$$

where  $\mathbf{E}(\cdot)$  is the expectation over an ensemble of independent runs,  $w_i$  are arbitrary positive weights such that  $\sum_i w_i = 1$ ,  $B^j$  is the cardinality of the support of  $j = \{m, \bar{p}_\mu\}$  and  $B^j \geq C^j \geq 1$ .

The use of a sufficiently large ensemble of runs allows to capture the overall degree of similarity between models and the data, washing away run-specific effects. Since ABMs might typically exhibit chaotic dynamics, stochastic shocks and/or tipping points, one run of the model might be completely different from the others.<sup>3</sup> To adequately explore behaviour of the model, a relatively large number of runs have to be considered and, when it comes to validation (after

<sup>3</sup>See the discussion in section 3.3 of Pyka and Fagiolo (2007) and, for a recent contribution on the issue of tipping points in macroeconomic agent based models, Gualdi et al. (2015).



having conveniently calibrated and/or estimated the model) different runs should ideally exhibit relatively similar dynamics. Finally, it is relevant to underline that, the application of (3) in the case of a deterministic model boils down exactly to (2).

Now, the only element that remains to be determined is the vector of aggregation weights,  $w_i$  with  $i = 1, \dots, L$ . In general, they are chosen to increase with  $l$  for two reasons. On one side, such choice reflects the greater importance assigned to patterns of similar behaviour lasting over longer time-windows and, on the other, it compensates for the increasing value of the logarithms' base  $a_l$ . A detailed discussion about weights' selection is found in Lamperti (2015, section 2.3), together with different robustness exercises showing results' poor sensibility to the choice of weights. For the purpose of this paper we consider *additively progressive* weights, that is, weights such that their first difference is constant and collectively sum up to one. This choice is additionally justified by the fact that additively progressive weights are unique; once  $L$  is fixed there is a unique vector satisfying previous requirements:

$$w_{l+1} = w_l + \frac{2}{L(L-1)} \quad \text{with } l = 1, \dots, L.$$

The *GSL-div* exhibits a set of interesting properties, even though it is important to recall it does not satisfy triangular inequality and, therefore, it is not a metric. In particular,

1. The *GSL-div* is well defined for all  $p$  and  $\bar{p}_\mu$
2.  $0 < GSL(p || \bar{p}_\mu) < 2$
3.  $GSL(p || \bar{p}_\mu) = H(p) \iff p = \bar{p}_\mu \quad \forall l = 1, \dots, L \text{ and } s \in S_{l,b}$ .

The first property guarantees that for any couple of probability vectors, the *GSL-div* between the two exists and can be computed independently of their support, which might either be the same or not.<sup>4</sup> The second property indicates that the *GSL-div* is bounded both from above and below. This is interesting and desirable for validation purposes, because benchmarks for extreme cases are naturally provided. Finally, property 3. shows the effective lower bound for the *GSL-div*: it is equal to the entropy of the real time series if and only if for every block length model  $\mu$  assigns each symbol the same probability the observed time series does. This would be the case of perfect matching between the dynamics of the simulated series and the data, whose patterns are mirrored exactly in all model's runs. Conversely, the *GSL-div* moves towards its upper bound, indicating that two series exhibit completely different behaviours, as soon as they tend to constantly persist over different states (i.e. symbols). For example, if we were to compare the dynamics of the inflation rates between two countries, this approach would detect maximally divergent behaviours when prices are constant in one country (zero inflation) but constantly rising (or falling) in the other or, in the same way, when deflation affects the first and hyperinflation the second.<sup>5</sup> However, it should be noticed that the upper bound of 2 constitutes a theoretical value, which is extremely unusual to be reached in practice. In Lamperti (2015) a more detailed set of properties is proved and discussed, including the link between the

<sup>4</sup>This constitutes a direct advantage vis-à-vis, for example, the Kullback-Leibler divergence (Kullback and Leibler, 1951).

<sup>5</sup>Note that in this example the observation units are inflation rates, and not price levels, which would deliver different results.

*GL-div* and the famous AIC (Akaike, 1973), which can be seen as a particular case of the present approach.

### 2.3 A simple example

Here I propose a simple and brief example showing how the *GSL-div* estimation works in practice. I consider three time series of length  $T = 10$ , called  $x$ ,  $y$  and  $z$  respectively, and omit dependence on time to ease notation. In particular,

$$\begin{aligned} x &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \\ y &= \{10, 9, 8, 7, 6, 5, 4, 3, 2, 1\} \\ z &= \{2, 2.5, 3, 3.5, 4, 4.5, 4, 4.5, 5, 5.5\}. \end{aligned}$$

These series might be thought as a real quantity (say,  $x$ ) and the output of two deterministic competing models ( $y$  and  $z$ ). Figure 1 graphically shows these three time series. By inspection, we notice that  $z$  behaves much more closely to  $x$  than  $y$  does; even though they have different slopes, they are both increasing over time, apart from one downward step in  $z$ . To the contrary,  $y$  is always decreasing, at a constant pace, which exactly corresponds to the opposite with respect to  $x$ 's one. Therefore,  $x$  and  $y$  touch the same states, each once and only once, but with reversed dynamics.

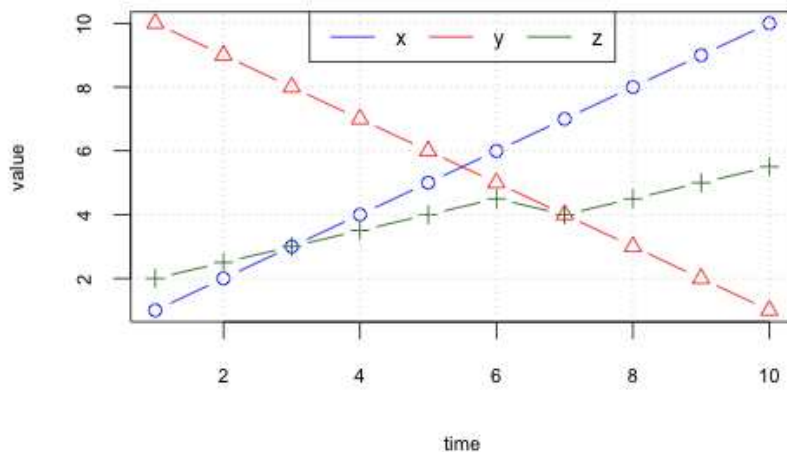


Figure 1: Three simple time series.

Now we want to employ the *GSL-div* to detect similarities between  $x$  and both  $y$  and  $z$ . Obviously, we expect it to identify a much closer dynamics between  $x$  and  $z$  rather than between  $x$  and  $y$ . First, we proceed with the symbolization process using  $b = L = 3$ . These values are chosen for exposition reasons; on one hand they are quite low, thereby making it more difficult for the *GL-div* to capture similarities in series' behaviour, on the other they allow for a relatively

short alphabet, which eases the representation of symbols. Table 1 reports symbols observed in the symbolized version of each series and their frequency. Each symbol represents a pattern that is observed in the data and its frequency measures the recursion of such symbol over time. The more similar frequency distributions over the different alphabets, the more time series exhibit analogous dynamics.

Table 1: Observed symbols and their frequencies for time series  $x$ ,  $y$  and  $z$ .

	$x$	$y$	$z$
length	observed symbols		
1	{1}, {2}, {3}	{1}, {2}, {3}	{1}, {2}, {3}
2	{11}, {12}, {22}, {23}, {33}	{11}, {21}, {22}, {32}, {33}	{11}, {12}, {22}, {23}, {32}, {33}
3	{111}, {112}, {122}, {222}, {223}, {233}, {333}	{111}, {211}, {221}, {222}, {223}, {233}, {333}	{111}, {112}, {122}, {223}, {232}, {233}, {323}, {333}
	frequencies		
1	0.4; 0.3; 0.3	0.4; 0.3; 0.3	0.3; 0.3; 0.4
2	0.33; 0.11; 0.22; 0.11; 0.22	0.33; 0.11; 0.22; 0.11; 0.23	0.22; 0.11; 0.11; 0.22; 0.11; 0.22
3	0.25; 0.125; 0.125; 0.125; 0.125; 0.125; 0.125;	0.25; 0.125; 0.125; 0.125; 0.125; 0.125; 0.125;	0.125; 0.125; 0.125; 0.125; 0.125; 0.125

Some similarity of behaviour between  $x$ ,  $y$  and  $z$  can be singled out just watching at the symbols reported in Table 1. Considering length equal to one ( $l = 1$ ), the focus falls on the persistence of each series within each state. Moving to  $l = 2$  patterns of length two are analysed. At this level it starts emerging the key difference between our three series: two are (almost always) increasing while the other decreases over time. This translates in the fact that supports of the distribution of time changes for  $x$  and  $z$  share a larger number of elements than it happens with  $y$ .<sup>6</sup> Beyond such similarities, the present approach recognizes from the very beginning the presence of a decreasing episode in  $z$ , which is absent in  $x$ . When blocks of length three are studied, the same evidences already emerged are confirmed and, in addition, we capture the fact that  $z$ 's downward phase lasts one period only.

The straightforward application of (2) leads then to the comparison of  $GSL-divs$  for the couples  $(x, y)$  and  $(x, z)$ . As expected,  $D_{GSL}(\mathbf{f}_x | \mathbf{f}_y) = 1.037 > 0.799 = D_{GSL}(\mathbf{f}_x | \mathbf{f}_z)$ , which gives some insight into the capacity of the present approach to identify similarities and differences in time series' dynamics even with low precisions ( $b$  and  $L$ ) and extremely short time series.

### 3 The Brock and Hommes Asset Pricing Model

To show the approach proposed in this paper more extensively, I rely on the widely known asset pricing model with heterogeneous agents proposed in Brock and Hommes (1998). The

<sup>6</sup>Some symbol is still present in three supports because of the low precision of the symbolization process, which does not allow to readily capture the strictly monotonic nature of series  $x$  and  $y$ . With  $b = 10$  for example,  $x$  and  $y$ 's supports at  $l = 2$  would be completely different.

model is ideal for illustrative purposes. It is relatively simple and costless to simulate but, on the other side, it offers variegated dynamics that are linked to a rich parameter space. Further, it has already been used as a test-model for different calibration and validation exercises (e.g. [Boswijk et al., 2007](#); [Recchioni et al., 2015](#)).

There is a population of  $N$  traders that can either invest in a risk free asset, which is perfectly elastically supplied at a gross return  $R = (1 + r) > 1$ , or in a risky one, which pays an uncertain dividend  $y$  and has a price denoted by  $p$ . Wealth dynamics is given by

$$W_{t+1} = RW_t + (p_{t+1} + y_{t+1} - Rp_t)z_t, \quad (4)$$

where  $p_{t+1}$  and  $y_{t+1}$  are random variables whose behaviour will be clarified in few lines and  $z_t$  is the number of the risky asset shares purchased at time  $t$ . In the market there is publicly available information on past prices and dividends, so that we can define the conditional expectation,  $E_t$ , and variance  $V_t$ . Traders are heterogeneous in terms of their expectations about future prices and dividends and are assumed to be myopic mean-variance maximizers. In particular, each agent demands a number of shares that solves

$$\max_{z_{h,t}} \left\{ E_{h,t}(W_{t+1}) - \frac{\alpha}{2} V_{h,t}(W_{t+1}) \right\}, \quad (5)$$

which implies

$$Z_{h,t} = E_{h,t}(p_{t+1} + y_{t+1} - Rp_t) / (\alpha\sigma^2), \quad (6)$$

where  $h$  denotes a trader-specific quantity,  $\alpha$  controls for the agents' risk aversion and  $\sigma$  indicates the conditional volatility, which is assumed to be equal across traders and constant in time. In the case of zero supply of outside shares and of different trader types, the market equilibrium equation can be written as

$$Rp_t = \sum n_{h,t} E_{h,t}(p_{t+1} + y_{t+1}), \quad (7)$$

where  $n_{h,t}$  denotes the share of type  $h$  traders at time  $t$ . In presence of homogeneous traders, perfect information and rational expectations it is possible to derive the following no-arbitrage market equilibrium condition:

$$Rp_t^* = E_t(p_{t+1}^* + y_{t+1}), \quad (8)$$

where the expectation is conditional on all histories of prices and dividends up to time  $t$  and where  $p^*$  indicates the fundamental price. In case the process of dividends is independent and identically distributed with time unvarying mean, equation (8) has a unique solution where the fundamental price is constant and such that  $p^* = E(y_t)/(R - 1)$ . In what follows it is convenient to express prices as their deviations from the fundamental,  $x_t = p_t - p_t^*$ .

Trading happens over a number of periods, denoted by  $t = \{1, 2, \dots, T\}$ . At the beginning of each trading period  $t$ , agents make expectations about future prices and dividends. I assume that agents are heterogeneous in that they have different forecasts of  $p_{t+1}$  and  $y_{t+1}$ . Beliefs

about future are assumed to take the following form:

$$E_{h,t}(p_{t+1} + y_{t+1}) = E_t(p_{t+1}^*) + f_h(x_{t-1}, \dots, x_{t-L}) \quad (9)$$

for all  $t$  and  $h$ . In words, investors believe that, in a heterogeneous world, prices may deviate from the fundamental value by some function  $f_h(\cdot)$  depending upon past deviations from the fundamental price. Many forecasting strategies have been implemented in the economic literature, specifying different trading behaviours and attitudes (Banerjee, 1992; Brock and Hommes, 1997; Lux and Marchesi, 2000; Chiarella et al., 2009). I follow Brock and Hommes (1998) in using a simple linear representation of beliefs:

$$f_{h,t} = g_h x_{t-1} + b_h \quad (10)$$

where  $g_h$  is said to be the trend component and  $b_h$  the bias of trader type  $h$ . If  $b_h \neq 0$ , we call agent  $h$  a pure trend chaser if  $g_h > 0$  (strong trend chaser if  $g > R$ ) and a contrarian if  $g < 0$  (strong contrarian if  $g < R$ ). If  $g_h \neq 0$ , type  $h$  is said to be purely biased (upward resp. downward biased if  $b_h > 0$  resp.  $b_h < 0$ ). The simple predictors in equation (10) could be considered as the simplest idealization of overreacting securities analysts or overreacting investors. In the special case  $g_h = b_h = 0$ , we obtain (pure) fundamentalists, who believes that prices return to their fundamental value. It is also possible to include a prototype of rational agent, who is characterized by  $f_{rational,t} = x_{t+1}$ . Rational agents have perfect foresight but, to obtain such a good prediction they are subjected to the payment of a cost  $C$ .<sup>7</sup>

To the purposes of the present application, I use a simple model with only two types of agents, whose behaviours vary according to the choice of trend components, biases and perfect forecasting costs. Combining equations (7), (9) and (10) it is possible to derive the following equilibrium condition:

$$Rx_t = n_{1,t}f_{1,t} + n_{2,t}f_{2,t}, \quad (11)$$

which allows to compute the price of the risky asset (in deviation from the fundamental) at time  $t$ . Traders' strategy is updated over time on the basis of accumulated wealth, which evolves according to equation (4). In particular the model allows for a switching behaviour that is governed by a parameter  $\beta$  in the following way. Each type  $h$  is associated with a fitness measure of the form:

$$U_{h,t} = (p_t + y_t - Rp_{t-1})z_{h,t} - C_h + \omega U_{h,t-1} \quad (12)$$

where  $\omega \in [0, 1]$  is a weight attributed to past profits. As time goes by, a strategy may become more profitable than the other one in term of fitness. All agent starts with their own (pre-specified) strategy, however at the beginning of each successive period they reassess the profitability of their own type relatively to others. The probability that a trader chooses the

---

<sup>7</sup>More in general, one could allow for the possibility that a positive a cost might be by paid also by non-rational traders; this is to mirror the fact that some trader might want to buy additional information which, however, might not be able to use (e.g. because of computational mistakes).

strategy  $h$  is given by “Gibbs” probability:

$$n_{h,t} = \frac{\exp(\beta U_{h,t})}{\sum_h \exp(\beta U_{h,t})}. \quad (13)$$

The rewind algorithm is designed so that the successful strategy gains a higher number of followers. In addition, algorithm introduces a certain amount of randomness, and more profitable strategies has a finite probability not to be preferred over less successful ones. In this way, the model capture imperfect information and bounded rationality of agents. This randomness also helps unlocking the system from the situation where all traders ends up with the same strategy  $h$ . The parameter  $\beta \in [0, +\infty)$  controls for the intensity of choice of the traders: the higher its values, the larger the likelihood of switching.

## 4 Model Selection and Validation

In this section I will illustrate and discuss the results obtained applying the *GSL-div* to the Brock and Hommes model described above, with two trader types denoted as 1 and 2. In particular, two main exercises are presented. First, I will show that the *GSL-div* is an adequate measure to distinguish between different versions and parameter configurations of the model. This is, for example, an explicit requirement in [Winker et al. \(2007\)](#). Secondly, I will move to the comparison of simulated dynamics, obtained through a calibrated version of the model, with real data from two major stock market indexes, namely the EuroSTOXX 50 and the CSI 300 (which represent the main European and Chinese markets respectively).

### 4.1 Discriminating among different models

To illustrate the ability of the *GSL-div* in distinguishing amongst different models, a known Data Generating Process (DGP hereafter) is needed, as it will be used as a benchmark against alternative and competing model configurations. Given the relatively high number of parameters in the model, a nearly infinite number of choices, delivering a wide and variegated array of dynamics (see [Brock and Hommes, 1998](#)) were available at this stage. To the purposes of the present exercise, the selection has tried to balance, on one side, the need to be clear and concise and, on the other, that of discriminating among objects that are inherently different but produce a relatively similar behaviour. Five models, in addition to the DGP, have been chosen and their configurations summarized in table 2. Despite the low number, they account for many different behaviours of the traders. The DGP is characterized by two trend-follower trader types, with the second extrapolating much stronger than the first, who is also upward biased while the other is not. The switching parameter ( $\beta$ ), which might take any positive value, is relatively low and in line with the numerical exercises carried out in [Brock and Hommes \(1998\)](#). The other models M1-M5 are obtained using the DGP as reference and modifying one or more characteristics defining the attitude of traders, while leaving unchanged those parameters that represents broader context conditions (e.g. the risk-free interest rate,  $r$ , or the volatility of the asset  $\sigma$ ) and summarized in the lower part of table 2. M1 and M2 modify the DGP in the intensity of choice of the two types, implying a much higher (M2) and lower (M1) likelihood of

switching towards the strategy delivering higher payoffs. M3 simply differs in the initial share of traders of first type, which is exactly balanced in the DGP while exhibits a strong dominance of second type in this model configuration. M4 maintains the bias of type 1 traders but assumes they are trend contrarians rather than followers, while type 2 keep their strategy but extract significantly more information from previous prices. Finally, M5 considers a fundamentalist trader vis-à-vis a trend follower, with the same switching attitude as modelled in the DGP. In addition to these features, it is relevant to point out that an element of randomness is included to enrich the framework and show the performance of the *GSL-div* in presence of noise and stochastic models. As in [Brock and Hommes \(1998\)](#), the dividend process,  $\{y_t\}_{t=1}^T$ , follows a stochastic process such that  $y_t = \bar{y} + \epsilon_t$  where the noise term  $\epsilon_t$  is i.i.d. uniformly distributed between -0.5 and +0.5 and, in our case,  $\bar{y} = 0$ . This formalization of the dividend process is kept unchanged for all the models considered.

Table 2: Parameters' value for Data Generating Process (DGP) and different models.

Parameter	Brief description	DGP	M1	M2	M3	M4	M5
$\beta$	intensity of choice	4	2	40	4	4	4
$n_1$	share of type 1 traders	0.5	0.5	0.5	0.1	0.5	0.5
$b_1$	bias of type 1 traders	0.2	0.2	0.2	0.2	0.2	0
$b_2$	bias of type 2 traders	0	0	0	0	0	0
$g_1$	trend component of type 1 traders	0.2	0.2	0.2	0.2	-0.2	0
$g_2$	trend component of type 2 traders	1.2	1.2	1.2	1.2	1.8	1.2
C	cost of obtaining type 1 forecasts	0	0	0	0	0	0
$\omega$	weight to past profits	0.5	0.5	0.5	0.5	0.5	0.5
$\sigma$	asset volatility	0.1	0.1	0.1	0.1	0.1	0.1
$\alpha$	attitude towards risk	10	10	10	10	10	10
r	risk-free return	0.1	0.1	0.1	0.1	0.1	0.1

*Note:* The symbol  $n_1$  is used to indicate the initial (at  $t = 0$ ) share of type 1 traders.

Figure 2 collects plots of a randomly chosen realization of the price process (in deviation from the fundamental value) produced both by the DGP (in red) and the various competing models (in blue). Direct inspection shows that notwithstanding each model accounts for different trading attitudes of the two types, the dynamics are quite similar, at least for an unaided eye. The relevant exception is provided by model M5, which robustly generates a continuously falling price. Whatever device, tool or methodology that aims at validating models, should be able to distinguish between those configurations yielding a truly different dynamics and, on the other side, to pull together those producing reasonably alike ones.

In our case, which treats an asset pricing model, dynamics are observed and compared with reference to two different quantities, that is, *prices* and *normalized returns*. More formally, if  $x_t$  is the price of a given asset at time  $t$  and  $\tau$  is the sampling frequency, the logarithmic difference of prices gives the returns:

$$r_t = \log(x_t) - \log(x_{t-\tau}) \approx \frac{x_t - x_{t-\tau}}{x_{t-\tau}}, \quad (14)$$

which can be normalized subtracting the longitudinal mean over the sample of interest



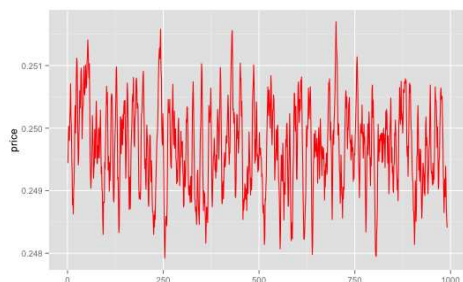
and dividing by the standard deviation,

$$nr_t = \frac{r_t - \langle r \rangle}{\sigma_r}, \quad (15)$$

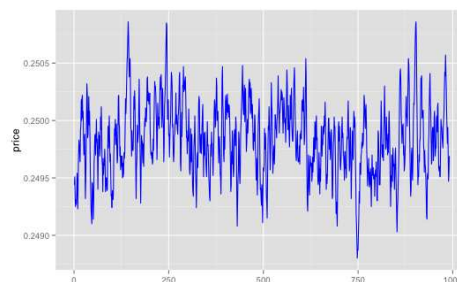
where  $\sigma_r$  is the standard deviation and  $\langle \cdot \rangle$  the time average over the considered period. Returns are normalized in many applications in order to symmetrize their distribution and to wash away the effects of their long run volatility.

The simulation setup is simple and constructed to mirror a real problem. The DGP is used to obtain a single realization, which is then labelled as the real world data. As it happens in practice, this series will be the unique term of comparison for all the five competing models and the DGP itself. I will test the ability of different models to replicate the dynamics observed in the data and I will rely only on the output of the simulations. Therefore, each of the configurations included in table 2 is initialized with the same conditions (but  $n_1$  in the case of M3) and used to

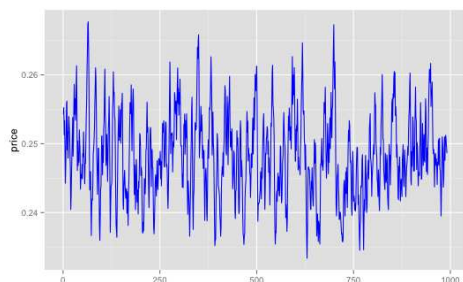
Figure 2: Randomly chosen realizations of Data Generating Process and other models.



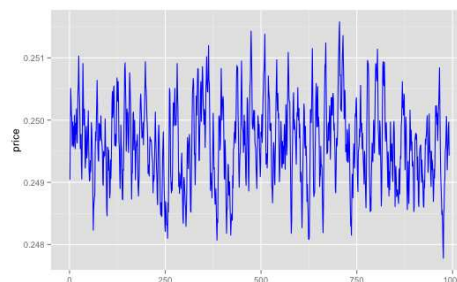
(a) Data Generating Process



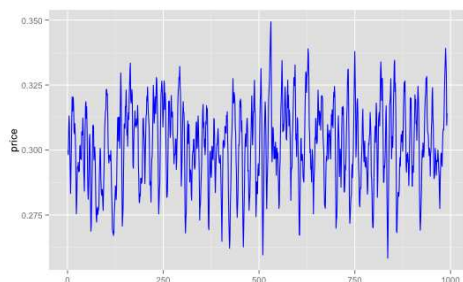
(b) Model 1



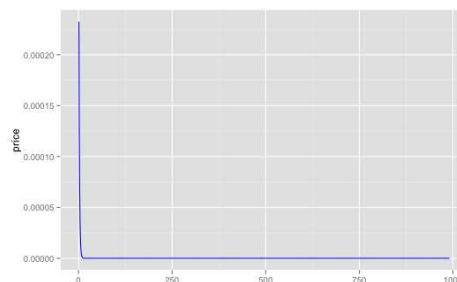
(c) Model 2



(d) Model 3



(e) Model 4



(f) Model 5



generate an ensemble of  $R = 500$  independent runs each of length  $T = 1000$ .<sup>8</sup> Then, the *GSL-div*, as expressed in equation (3), is employed to assess the similarity between what we have called real world data and models output. It is relevant to remark that the two free parameters in the procedure leading to the estimation of the *GSL-div*, that is, the precision of the symbolization ( $b$ ) and the maximum blocks' length ( $L$ ) are set accordingly to Lamperti (2015).<sup>9</sup> Table 3 reports both the distances estimated for different lengths of symbols and the *GSL-div*.

Table 3: *GSL-div* between data and models for price and normalized returns.

Prices							
block length	weights	DGP	M1	M2	M3	M4	M5
1	0.05	0.739640	1.000337	0.859058	0.739736	0.741820	1.565000
2	0.10	0.610180	0.764265	0.681810	0.610093	0.673059	1.031640
3	0.14	0.565688	0.683830	0.622171	0.565782	0.590420	0.854371
4	0.19	0.542423	0.646385	0.592035	0.542455	0.558736	0.774315
5	0.24	0.528241	0.634290	0.573849	0.528183	0.539089	0.747396
6	0.29	0.516632	0.631058	0.559819	0.516771	0.526844	0.750571
<b>GSL-div</b>		<b>0.550845</b>	<b>0.672557</b>	<b>0.604071</b>	<b>0.550887</b>	<b>0.569079</b>	<b>0.834717</b>
<b>MC s.d.</b>		<b>0.008300</b>	<b>0.013780</b>	<b>0.014007</b>	<b>0.008000</b>	<b>0.006480</b>	<b>0.000000</b>
Normalized Returns							
block length	weights	DGP	M1	M2	M3	M4	M5
1	0.05	0.934609	1.042379	0.971981	0.934126	0.980297	1.544845
2	0.10	0.921512	1.011694	0.951396	0.921089	0.966483	1.294408
3	0.14	0.916102	0.996014	0.944651	0.915776	0.959992	1.178728
4	0.19	0.908012	0.969536	0.932853	0.907875	0.954991	1.173793
5	0.24	0.879335	0.925125	0.902374	0.879559	0.939477	1.262931
6	0.29	0.810996	0.887263	0.849655	0.812325	0.903309	1.267832
<b>GSL-div</b>		<b>0.877173</b>	<b>0.946722</b>	<b>0.907140</b>	<b>0.877470</b>	<b>0.939545</b>	<b>1.251746</b>
<b>MC s.d.</b>		<b>0.009960</b>	<b>0.027033</b>	<b>0.009171</b>	<b>0.009290</b>	<b>0.010230</b>	<b>0.006670</b>

As a first observation, one can notice that the *GSL-div* distinguishes clearly among the majority of proposed models, both for what concerns the dynamics of prices and those of returns, delivering consistent results. In particular, the DGP is correctly identified as the closest model to the real world data, and the low Monte Carlo standard deviation allows to conclude that its distance is statistically different from those found for any other model but one (M3).<sup>10</sup> This result is extremely reasonable and helps prove the good performance of the *GSL-div*. In particular, M3 exhibits the same parametric structure of the DGP and differs only in the initial

<sup>8</sup>The unique exception is the DGP, which is run  $R + 1$  times. Then one of these realizations is randomly selected as the real data while the others  $R$  are used to compare the DGP with these data.

<sup>9</sup>Therefore,  $b = 5$  and  $L = 6$ . In Lamperti (2015) the *GSL-div* is proved to be robust to changes in these two parameters. Their choice is problem-specific and depends on the degree of precision requested by the modeller, the number of competing models or configurations, the time scale of interest and the available computational power. In practical situations, as a rule of thumb, one can select the combination of  $b$  and  $L$  starting by  $b = L = 2$  and then increasing their value one at the time stopping when any further increase in one and the other parameter do no change the order of models provided by the *GSL-div*. That will be the minimum order or complexity which is needed to discriminate robustly among a set of models. In my experience,  $b = 5$  and  $L = 6$  are generally sufficient to the scope.

<sup>10</sup>The Monte Carlo standard deviation is simply the standard deviation of the *GSL-div* computed in (2) across the ensemble. Being the latter composed by 500 independent runs of the same model, it is possible to interpret it as a Monte Carlo exercise on the seed of the random number generator

number of type 1 traders. However, being  $n_1 = 0.1$  a dis-equilibrium starting point,  $\beta = 4$  a reasonably high intensity of choice and having M3 the same stable steady state as DGP (see [Brock and Hommes, 1998](#), for details on steady states of the model with two types), the share of type 1 traders suddenly converges to the equilibrium value and then fluctuates around it because of the effects introduced by the random noise. From this point on, M3 and DGP can be seen as completely identical models and, therefore, it is a good signal that the *GSL-div* cannot distinguish significantly between the two. Finally, as a consistency check, it is worth remarking that M5, which exhibits a dynamics of prices clearly at odds with the real data, is successfully found as the most distant model.

## 4.2 Validation against real data

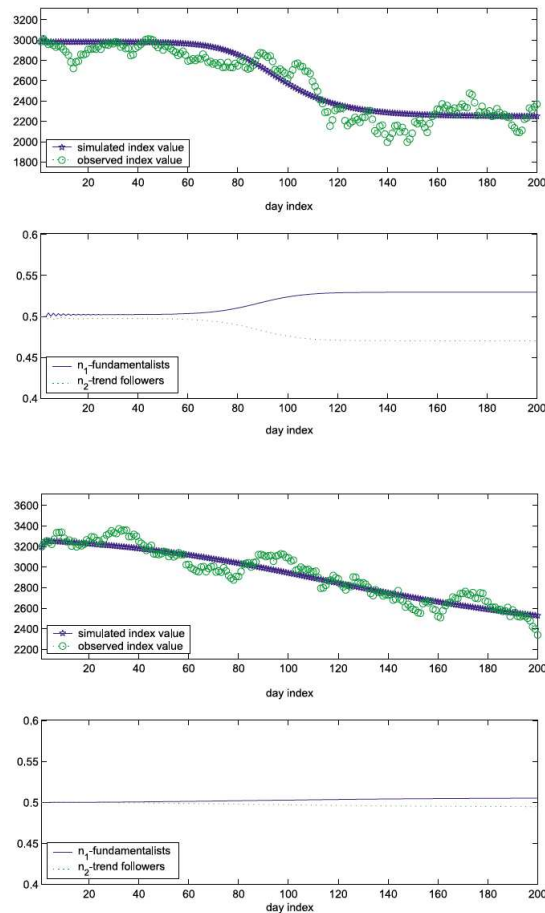
Once the *GSL-div* has been proven to successfully discriminate amongst different models, it can be used to validate them against actual data. To this purpose I draw on the results obtained in [Recchioni et al. \(2015\)](#), where a particular version of the Brock and Hommes model described in section 3 has been estimated using real stock market data. Specifically, this simplified model does not include dividends and has been constrained to incorporate two particular types of agents, namely a pure fundamentalist ( $g_1 = b_1 = 0$ ) and an unbiased trend follower ( $g_2 > 0$  and  $b_2 = 0$ ), and it has been calibrated using daily data covering the period ranging from February 25, 2011 to December 16, 2011, for a total of 200 observations. In what follows, I refer exactly to this interval and target two major stock market indexes, namely the EuroSTOXX 50 (which is an index composed by main corporations in the Euro area) and the CSI 300 (which is one of the most important Asian indexes, designed to replicate the performance of 300 stocks traded in the Shanghai and Shenzhen stock exchanges). Table 4 collects the values of parameters set or calibrated in [Recchioni et al. \(2015\)](#). In the exercises that follows some of these parameters will be used as benchmark. Figure 3 shows the behaviour produced by the model calibrated

Table 4: Parameters of the calibrated model for EuroSTOXX 50 and CSI 300.

	$\beta$	$g_1$	$g_2$	$\alpha$	$p^*$	C	$\omega$	$\sigma$
EuroSTOXX 50	0.642	0	2.0	18.207	0.746	0	1	0.1
CSI 300	0.078	0	1.996	13.999	0.682	0	1	0.1

on the EuroSTOXX (top two panels) and on the CSI (bottom two panels), together with the real data and the share of traders following one or the other strategy along the simulation. It is evident that, despite being a simple model, it provides a reasonably good performance in tracking the real indexes. The additional value brought by the use of ABMs should be the possibility to analyse the micro-determinants of these macro-behaviours and, for example in this case, to single out differences in the attitude of traders operating in different geographical areas. In what follows I perform two different exercises by means of the *GSL-div*. First, I explore the similarity between real data and model configurations obtained maintaining the same structure (fundamentalists vs. trend followers) and parameter values as in table 4, but I let the trend-following component,  $g_2$ , and the switching parameter,  $\beta$ , to considerably vary. Second, I allow the model to account for richer combinations of traders' attitudes (e.g. two trend

Figure 3: Actual prices and model simulation in the calibration interval. Source: [Recchioni et al. \(2015\)](#).



*Note:* This figure is composed by four panels. The first two from the top refers to the EuroSTOXX index and the corresponding model, while the bottom two to the CSI. For each index, the first (top) panel plots the behaviour of the price (observed and simulated) in the time interval used for calibration, while the second (bottom) the share of the two types of traders in the same time interval.

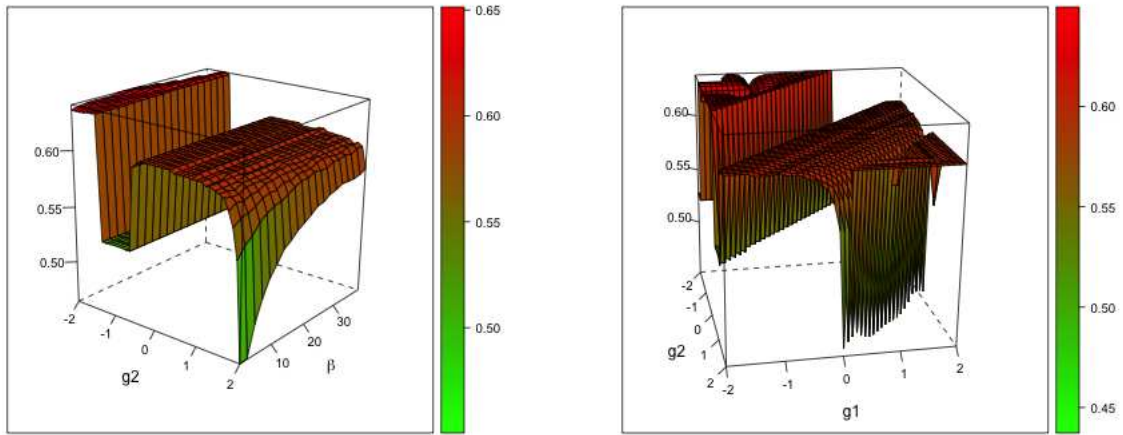
followers, one trend-follower and one contrarian, one contrarian and one fundamentalist) and I check whether some of them are able to provide a better account for the dynamics observed in the data. Obviously enough, if the calibrated model turns out to be the closest to the data both for EuroSTOXX and CSI indexes and, keeping fixed its structure, no combinations of relevant parameters (e.g the intensity of choice) produce significantly better results, I will conclude in favor of the empirical validity of the model, at least for what concerns its ability to track the historical behaviour of the targeted system.<sup>11</sup>

To start with, I need to build a convenient subspace of parameters whose points will be used to construct model configurations that, in turns, will serve to produce simulated output to feed the *GSL-div*. In the case of the first exercise, I consider all possible combinations of parameters' values found in a two-dimensional grid obtained using the following intervals  $1 \leq \beta \leq 40$  and  $-2 \leq g_2 \leq 2$ , where the former is discretized in 21 equally spaced segments, while the latter in

<sup>11</sup>This paper only presents an illustration of the features, performance and exercises that is possible to obtain using the *GSL-div*. A more complete analysis would be needed to assess exhaustively the empirical validity of the model.

41. For the second exercise, instead, the same interval used for  $g_2$  is allowed to characterize the trend component of the first trader type.<sup>12</sup> This procedure leads to the inclusion of 861 model configurations for the first exercise and 1681 for the second. All these models are run starting from the same initial conditions and for  $T = 200$  periods. The *GSL-div* is used to measure their distance with respect to the real data (both EuroSTOXX and CSI). Figure 4 shows the results obtained for the first (a) and second (b) exercise using EuroSTOXX data, while figure 5 does the same for the CSI index.

Figure 4: *GSL-div* between model and EuroSTOXX50 for different portions of the parameters space.



(a) Space generated by  $\beta$  and  $g_2$  ( $g_1 = 0$ ).

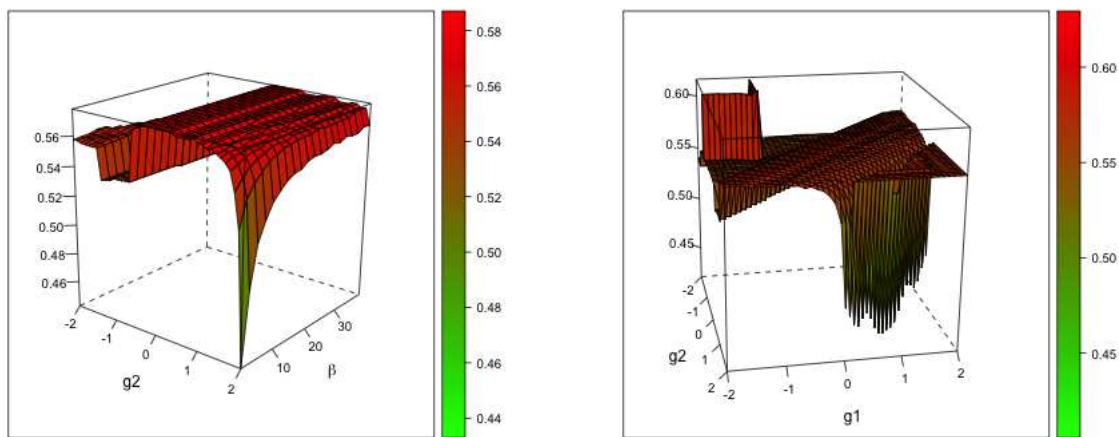
(b) Space generated by  $g_1$  and  $g_2$  ( $\beta = 0.642$ ).

Different remarks apply. To begin with, subfigures 4a and 5a contains two insights. On one hand they show that, in large parts of the explored parameters' subspace, the dynamics of the Brock and Hommes model with a trend follower and a fundamentalist types are guided by  $g_2$ , that is, by the strength at which the trend followers extrapolate information on the basis of past observations. This finding is partially in line with Teräsvirta (1994), Boswijk et al. (2007) and Recchioni et al. (2015), where it is shown that the intensity of choice has little significance in switching models. However, a notable difference emerges:  $\beta$  seems not to affect the distance between models and data for the majority of values, but in those cases where it is much higher than data would admit, having the correct specification of the trend-following attitude does not suffice to obtain behaviours consistent with the data. Moreover, the Brock and Hommes model is found to be as much sensitive to changes in the switching parameter with respect to the best values as it is to  $g_2$ , as confirmed by the steepness of the *GSL-div* function represented in subfigures 4a and 5a. Such sensitivity appears stronger when the model is calibrated to CSI data. Pushing forward these reasoning one can notice that the model is deeply affected by tipping points in the space of parameters, that is areas where the behaviour of the model changes suddenly and dramatically in response to small variations in conditions (see also Gualdi et al., 2015). For example, in the case of the EuroSTOXX calibration, subfigure 4a shows a not so small area where  $g_2$  is around -1.5 and the model shifts immediately its behaviour in response to

<sup>12</sup>These intervals allow to encompass many values of relevant parameters that are treated in the literature (see Brock and Hommes, 1998; Boswijk et al., 2007).

a small change in the strength of the trend chasers attitude and irrespectively of  $\beta$ ; then it keeps constant for a while (relatively to the real data) and shifts again. A second remark concerns the general behaviour of the model calibrated in [Recchioni et al. \(2015\)](#). Surfaces constructed through the *GSL-div* confirm that a switching parameter below 3 coupled with a trend-following intensity really close to 2 deliver, by far, the most similar dynamics with respect to the data in both the two analysed markets. Finally, comparing the EuroSTOXX and CSI cases one can notice that the role of the switching parameter is much more important in the Asian market rather than in the European one, meaning that the latter is more compatible, in relative terms, with a larger set of values for  $\beta$ . This is rapidly explained by the different behaviour of the data, where it is possible to see that, in the considered period, CSI exhibits an approximately linear negative trend while the EuroSTOXX does not. Since the switching parameters controls, other things being equal, for the steepness of the downward dynamics (larger  $\beta$  corresponds to an s-shaped behaviour with increasing steep of the s), it is natural that a linear trend is compatible only with a much more restricted portion of the parameter space.

Figure 5: *GSL-div* between model and CSI300 for different portions of the parameters space.



(a) Space generated by  $\beta$  and  $g_2$  ( $g_1 = 0$ ).

(b) Space generated by  $g_1$  and  $g_2$  ( $\beta = 0.078$ ).

Even though the calibrated Brock and Hommes model accounting for trend followers and a pure fundamentalists is decently able to replicate the dynamics observed in the real data, much more caution is suggested by subfigures 4b and 5b. They indicate that, keeping fixed other parameters, there are many combinations of traders attitudes that guarantee approximately the same behaviour with respect to the data. This finding is robust across the different stock markets considered and, in a comparative perspective, it is more evident for the EuroSTOXX case, where there are two separate areas delivering a good matching with the data. This simple problem of multiple minima in the subspace of parameters spanned by  $g_1$  and  $g_2$  might be particularly harmful for the empirical validity of the model, because it supports the claim that many different combinations of traders' attitudes are compatible with the same dynamics observed in the data. In addition, since the estimated *GSL-divs* for these model configurations are pretty much close one to the other, it hardly arguable why one (e.g. involving trend followers and fundamentalists) should be preferred to another (e.g. involving two trend follower types

with different extrapolation strengths) in the attempt to explain some dynamics of the price.

To better investigate this issue I enlarged the possible combinations of traders' attitudes by considering a new grid where  $g_1$  and  $g_2$  are allowed to take values ranging from  $-4$  (which models a really strong trend contrarian type) and  $+4$  (which models a really strong trend follower type). As a result, 861 additional model configurations are tested against the EuroSTOXX and CSI by means of the *GSL-div*. Table 5 reports the configurations yielding the lowest ten values of the *GSL-div*.

Table 5: Model configurations yielding the 10 lowest *GSL-div* values.

EuroSTOXX 50			CSI 300			
rank	$g_1$	$g_2$	<i>GSL-div</i>	$g_1$	$g_2$	<i>GSL-div</i>
1	1.2	0.8	0.450575	1.2	0.8	0.394348
2	1.4	0.6	0.450963	1	1	0.395828
3	1	1	0.451525	1.4	0.6	0.400041
4	0	2	0.457319	1.6	0.4	0.414634
5	1.6	0.4	0.457503	0	2	0.418405
6	1.8	0.2	0.461868	1.8	1.2	0.424130
7	2.2	-0.2	0.462625	2.2	-0.2	0.425081
8	2.4	-0.4	0.476489	2.4	-0.4	0.440627
9	2.6	-0.6	0.491565	2.6	-0.6	0.456078
10	2.8	-0.8	0.504789	2.8	-0.8	0.470646

Variegated combinations of the behavioural strategies adopted by the two traders types are found within the ten cases that are closest to the data, and the configuration assuming a pure fundamentalist and a trend follower does not provide the best result in any of the two markets analysed in this paper. What emerges as a general trait is that the sum of the trend components for the two agents has to be close to 2 in order for the model to be consistent with the dynamics in the data. While the same happens considering parameters estimated in [Boswijk et al. \(2007\)](#) and [Recchioni et al. \(2015\)](#) (confirming again the *GSL-div* provide results consistent with other methodologies), it is now showed that what matters is not the strict presence of mean-reverting and trend follower types (as in [Boswijk et al., 2007](#)). Rather, the model has to account for a strong trend following component, which might either come from a unique type that heavily extrapolates information from past observations or the combinations of different types with milder, or even opposite, attitudes towards the trend. However, results suggest that if one of the traders types follows the trend too strongly, a compensating trend contrarian type reduces the bullish pressure created in the market. Summing up, this paper uses the *GSL-div* to explore and validate the Brock and Hommes asset pricing model and finds it reasonably able to resemble dynamics observed in actual stock markets, provided that it accounts for a balanced trend following attitude, which destabilizes the asset market from the fundamentals but not so strongly to create large bubbles (with price diverging to  $+\infty$ ).



## 5 Conclusions

Validation of simulated models is still an open issue. One way of tackling this problem is via the identification of a measure quantifying the distance between simulated and real-world data with respect to the observed dynamics. This paper presents an illustrative application on the use of the *GSL-div* developed in [Lamperti \(2015\)](#) to the validation of simulated models. In particular, different versions of the asset pricing model proposed in [Brock and Hommes \(1998\)](#) are analysed. The proposed approach is found to successfully discriminate amongst alternative, competing models with reasonable precision, both when price or return dynamics are at stake. The Brock and Hommes model calibrated in [Recchioni et al. \(2015\)](#) is then validated against data from two major stock market indexes, namely the EuroSTOXX 50 and the CSI 300. What emerges is that when the model is constrained to include two specific types of traders (namely a fundamentalist and a trend follower) it seems to achieve a reasonable similarity with the real data; however, when different combinations of traders' types are allowed, it is difficult to argue what configuration should be preferred over another, since many different attitudes are compatible with the same dynamics as those observed in actual data. However, what emerges as a general trait, is that empirical validity of model, requires to account for a strong trend following component, which might come both from a unique trend follower type that heavily extrapolates information from past observations or the combinations of different, milder, or even opposite, trend follower types. Finally, it is worth to recall that even though in this paper the *GSL-div* is used as a validation tool applied to already calibrated models, an interesting application will be the development of a calibration procedure including it in the objective function.

## Acknowledgements

The author would like to thank Mattia Guerini, Mauro Napoletano and Andrea Roventini for valuable comments and suggestions. All the shortcomings are the author's.

## References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Kiado, A., editor, *Proceedings of the Second International Symposium on Information Theory*, pages 267–281, Budapest.
- Alfarano, S., Lux, T., and Wagner, F. (2005). Estimation of agent-based models: the case of an asymmetric herding model. *Computational Economics*, 26(1):19–49.
- Alfarano, S., Lux, T., and Wagner, F. (2006). Estimation of a simple agent-based model of financial markets: An application to australian stock and foreign exchange data. *Physica A: Statistical Mechanics and its Applications*, 370(1):38–42.
- Amilon, H. (2008). Estimation of an adaptive stock market model with heterogeneous agents. *Journal of Empirical Finance*, 15(2):342 – 362.

- Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.
- Barde, S. (2015). A practical, universal, information criterion over nth order markov processes. Kent Discussion Paper 15/04.
- Bianchi, C., Cirillo, P., Gallegati, M., and Vagliasindi, P. A. (2007). Validating and calibrating agent-based models: a case study. *Computational Economics*, 30(3):245–264.
- Bianchi, C., Cirillo, P., Gallegati, M., and Vagliasindi, P. A. (2008). Validation in agent-based models: An investigation on the cats model. *Journal of Economic Behavior & Organization*, 67(3):947–964.
- Boswijk, H. P., Hommes, C. H., and Manzan, S. (2007). Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control*, 31(6):1938 – 1970. Tenth Workshop on Economic Heterogeneous Interacting Agents WEHIA 2005.
- Brock, W. A. and Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65(5):1059–1095.
- Brock, W. A. and Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22(8):1235 – 1274.
- Canova, F. and Sala, L. (2009). Back to square one: identification issues in dsge models. *Journal of Monetary Economics*, 56(4):431–449.
- 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.
- Dawid, H. and Fagiolo, G. (2008). Agent-based models for economic policy design: Introduction to the special issue. *Journal of Economic Behavior & Organization*, 67(2):351–354.
- Dosi, G., Fagiolo, G., Napoletano, M., and Roventini, A. (2013). Income distribution, credit and fiscal policies in an agent-based keynesian model. *Journal of Economic Dynamics and Control*, 37(8):1598–1625.
- Dosi, G., Fagiolo, G., Napoletano, M., Roventini, A., and Treibich, T. (2015). Fiscal and monetary policies in complex evolving economies. *Journal of Economic Dynamics and Control*, 52(0):166 – 189.
- Dosi, G., Fagiolo, G., and Roventini, A. (2010). Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767.
- Fabretti, A. (2012). On the problem of calibrating an agent based model for financial markets. *Journal of Economic Interaction and Coordination*, 8(2):277–293.
- Gallegati, M. and Richiardi, M. (2009). *Agent based modelling in economics and complexity*, volume In: Meyer, R.A. (Ed.), *Encyclopedia of Complexity and Sistem Science*. Springer, New York, USA.



- Giannone, D., Reichlin, L., and Sala, L. (2006). Vars, common factors and the empirical validation of equilibrium business cycle models. *Journal of Econometrics*, 132(1):257–279.
- Gilli, M. and Winker, P. (2003). A global optimization heuristic for estimating agent based models. *Computational Statistics & Data Analysis*, 42(3):299–312.
- Gourieoux, C. and Monfort, A. (1997). *Simulation based econometric methods*. Oxford University Press.
- 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.
- Grazzini, J., Richiardi, M., and Tsionase, M. (2015). Bayesian estimation of agent-based models. Technical report, Laboratorio R. Revelli, Centre for Employment Studies.
- Gualdi, S., Tarzia, M., Zamponi, F., and Bouchaud, J.-P. (2015). Tipping points in macroeconomic agent-based models. *Journal of Economic Dynamics and Control*, 50:29 – 61. Crises and Complexity Complexity Research Initiative for Systemic Instabilities (CRISIS) Workshop 2013.
- Guerini, M. and Moneta, A. (2016). A method for agent-based model validation. LEM Papers Series 2016/16, Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies, Pisa, Italy.
- Jacob Leal, S., Napoletano, M., Roventini, A., and Fagiolo, G. (2015). Rock around the clock: An agent-based model of low- and high-frequency trading. *Journal of Evolutionary Economics*, 26(1):49–76.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22:49–86.
- Lamperti, F. (2015). An Information Theoretic Criterion for Empirical Validation of Time Series Models. LEM Papers Series 2015/02, Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies, Pisa, Italy.
- LeBaron, B. and Winker, P. (2008). Introduction to the Special Issue on Agent-Based Models for Economic Policy Advice. *Journal of Economics and Statistics*, 228(2+3):141–148.
- Leombruni, R., Richiardi, M., Saam, N. J., and Sonnessa, M. (2006). A common protocol for agent-based social simulation. *Journal of Artificial Societies and Social Simulation*, 9(1):15.
- Lin, J. (1991). Divergence measures based on the shannon entropy. *IEEE Transactions on Information theory*, 37:145–151.
- 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.
- Manson, S. (2002). *Validation and verification of multi-agent systems, in Complexity and Ecosystem Management*. Cheltenham: Edward Elgar, edited by m.a. janssen edition.

- Marks, R. (2013). Validation and model selection: Three similarity measures compared. *Complexity Economics*, 2(1):41–61.
- Paccagnini, A. (2009). Model validation in the dsge approach: A survey. Working Paper, mimeo.
- Pellizzari, P. and Forno, A. D. (2006). A comparison of different trading protocols in an agent-based market. *Journal of Economic Interaction and Coordination*, 2(1):27–43.
- Pyka, A. and Fagiolo, G. (2007). Agent-based modelling: a methodology for neo-schumpeterian economics. In Pyka and Horst, editors, *Elgar companion to neo-schumpeterian economics*. Edward Elgar Publishing.
- Recchioni, M. C., Tedeschi, G., and Gallegati, M. (2015). A calibration procedure for analyzing stock price dynamics in an agent-based framework. *Journal of Economic Dynamics and Control*, 60:1 – 25.
- Rosen, R. (1985). *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*. Oxford: Pergamon.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218.
- Windrum, P., Fagiolo, G., and Moneta, A. (2007). Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, 10(2):8.
- 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(2):125–145.