Estimation and Validation of Agent-based Computational Economic Models

Guerini Mattia

October 30, 2020

Université Côte d'Azur - GREDEG

Table of Contents

Critique to ACE models

- *ad-hoc* behavioural assumptions;
- large numbers of free parameters – somehow too many;
- simulations with some incredible values.

The last two are **technical and valuable critiques** (acknowledged also by AB modellers)

However:

- many parameters in ABM are *deep* and can therefore be estimated;
- there exist tools to estimate ACE models;
- ABM are not the unique ones facing these problems (Fagiolo and Roventini, 2017).

Glossary and definitions

Demonstration models
Existence proofs for phenomena of interest. They are also called *exploration models* because of their primary objective is that of unveiling the presence of some possible mechanism at work.

Descriptive models
Attempt to track dynamic historical phenomena. Their primary aims are that of qualitatively or/and quantitatively mimic the dynamics of some real phenomenon and to explain it by means of the economic forces embedded in the model.

Both types of models require verification; only the second class requires validation (and credible values).

Stylized fact (SF)

An empirical/statistical regularity.

> *Facts as recorded by statisticians are always subject to numerous snags and qualifications, and for that reason are incapable of being summarized.*
>
> *Kaldor (1957)*

SF can be characterized by different degrees of:

- robustness;
- generality;
- easiness of generation.

Replication of SF does not mean the model is a descriptive one, but a descriptive model shall replicate (at least some) SF (Brock, 1999; Buchanan, 2012; Fagiolo et al., 2019; Guerini and Moneta, 2017).

Parameters' space exploration (*sensitivity analysis*)
Replies to the question: *How the model's outcome is affected by a variations in one parameter?*
Is a sort of model elasticity w.r.t. a change in the parameters (Salle and Yıldızoğlu, 2014).

Trade-off:

- **robust model**: ideally you don't want that all the results change as soon as you change one parameter (unless that's the scope of the parameter within the model);
- **tautological model**: ideally you don't want a model whose outcome is equivalent for all the values of a parameter (in that case it is likely that the parameter does not play any role and can be dropped).[1]

[1] Under the assumption that a parsimonious model is better than an overly complicated one.

Model stationarity and model ergodicity

Model Stationarity

A model is said to be stationary if, after a certain period T_s, the main variable(s) under scrutiny converges to a time-invariant distribution;

Model Ergodicity

A model is said to be ergodic if, after a certain period T_s, the main variable(s) under scrutiny converges to a time-invariant distribution and the distribution is the same for all the Monte Carlo simulations.

Verification and validation (V&V)

Model Verification

Replies to the question *are we building the model right?*

It is a formal requirement on the efficacy of the code to be correctly representing the conceptual model under study (i.e. the economic forces at stake) (Yilmaz, 2006; Naylor and Finger, 1967). Any model should satisfy this requirement.

Model Validation

Replies to the question *are we building the right model?*

It consists of measuring and evaluating the extent to which the model that we have designed is a good representation of the reality (Marks, 2013). Not all models are built to be validated.

Model Calibration
Finding the parameter values that allow the model to resemble as close as possible some properties of the data (ex-ante, it does not pretend that the model is a representation of the true DGP).

Model Estimation
Finding the parameters values that are most likely to have generated the observed empirical data series (ex-ante, it assumes that the model is the true representation of the DGP).

Note: there is only a philosophical difference between the two. Methodologically, they converge to overlapping exercises.[2]

[2] In what follows I will use the two concepts as identical.

Estimation of ACE models

Indirect inference

A simulation-based method for estimating the parameters of economic models where the likelihood function is intractable (Gourieroux et al., 1993).

General procedure

- simulate data from the economic model for different values of its parameters;
- use an *auxiliary model* to form a criterion function;
- determine the parametrization at which the model and the observed data look most similar.

The objective is that of choosing the parameter vector such that the observed data and the simulated data look most similar from the point of view of the chosen auxiliary model.

The ACE model

Let:

$$y_t = \mathcal{G}(y_{t-1}, x_t, u_t; \theta), \quad t = 1, \dots, T, \quad y_0 = \hat{y}_0, \quad u_t \sim iid \tag{1}$$

describe the DGP (i.e. our ACE model). $\mathcal{G}$ can be interpreted as a probability density function for y_t conditional on y_{t-1} and x_t, depending also on the k-dimensional parameter vector θ.

An econometrician would approach the problem by using the observed data to estimate the k-dimensional parameter vector $\hat{\theta}$.

Standard inference tools – e.g. Maximum Likelihood Estimation (MLE) – require that $\mathcal{G}$ has an analytical closed form such that a likelihood function can be derived and the optimal θ^* can be estimated.

Indirect inference helps when $\mathcal{G}$ is too complex to have a closed form. That's why it is particularly suited for ABM, where $\mathcal{G}$ typically represents a high-dimensional and possibly non-linear model.

The auxiliary model (on actual data)
Let:

$$\mathcal{F}(y_t|y_{t-1}, x_t; \beta) \tag{2}$$

be an auxiliary model (sufficiently simple) depending on the p-dimensional parameter vector β (with $p \geq k$) which can also be incorrectly specified.

The parameters $\hat{\beta}$ can be estimated using the actual data as:

$$\hat{\beta} = \underset{\beta}{argmax} \sum_{t=1}^{T} log\mathcal{F}(y_t|y_{t-1}, x_t; \beta) \tag{3}$$

The auxiliary model (on simulated data)
The parameters β of the auxiliary model $\mathcal{F}$ can also be estimated over the simulated data.

Procedure:

- draw the disturbances $u_{m,t}$, with $m = 1, \ldots, M$ and $t = 1, \ldots, T$ from a defined probability distribution;
- pick a parameter vector θ and iterate the model $\mathcal{G}(\theta)$ to generate a random sequence of the variables of interest $\tilde{y}_{m,t}(\theta)$;
- maximize the average of the log-likelihood across the M simulations to obtain:

$$\tilde{\beta}(\theta) = \underset{\beta}{argmax} \sum_{m=1}^{M} \sum_{t=1}^{T} log\mathcal{F}(\tilde{y}_{m,t}(\theta) | \tilde{y}_{m,t-1}(\theta), x_t; \beta) \ ; \tag{4}$$

- select the values θ that minimizes the distance between $\hat{\beta}$ and $\tilde{\beta}(\theta)$.

When using indirect inference, it is necessary to choose a metric for measuring the distance between the parameters of the auxiliary model estimated respectively with real and simulated data.

There are three possibilities corresponding to the classical trinity of testing:

$$\hat{\theta}^{WA} = \underset{\theta}{argmin}\left[\left(\hat{\beta} - \tilde{\beta}(\theta)\right)' W\left(\hat{\beta} - \tilde{\beta}(\theta)\right)\right]$$

$$\hat{\theta}^{LR} = \underset{\theta}{argmin}\left[\sum_{t=1}^{T} log\mathcal{F}(y_t|y_{t-1}, x_t; \hat{\beta}) - \sum_{t=1}^{T} log\mathcal{F}(y_t|y_{t-1}, x_t; \tilde{\beta}(\theta))\right]$$

$$\hat{\theta}^{LM} = \underset{\theta}{argmin}\left[S(\theta)' V S(\theta)\right], \qquad S(\theta) = \sum_{m}\sum_{t} \frac{\partial}{\partial \beta}\left[log\mathcal{F}\left(\tilde{y}_t^m(\theta)|\tilde{y}_{t-1}^m, x_t; \hat{\beta}\right)\right]$$

where W and V are two positive-definite weighting matrices.

Pros and cons for ACE applications

Advantages

- relies on estimators with well known statistical properties;
- even if the auxiliary model is incorrect, the estimation can be correct;
- it is the natural extension of a classical econometric problem.

Drawbacks

- the estimation error increases with the dimensionality of the problem;
- requires the simulation of the true model over a large set of parameter values and with many Monte Carlo runs (might be computationally heavy);
- results on the optimal choice of θ might depend upon:
  - the chosen objective variables;
  - the chosen distance measure;
  - the number of Monte Carlo replications.

Particular cases

The **Simulated Minimum Distance** (SMD - Hsiao, 1989) and the **Method of Simulated Moments** (MSM - McFadden, 1989) can be seen as special cases of the indirect inference, where the auxiliary models on which the estimation exercise is based are:

- one or few parameters estimated from a simple econometric model involving the variables of interest (e.g. the Pareto-tail parameter of the size distribution of firms);
- one or few low order moments concerning the variables of interest (e.g. average growth rate of GDP).

Estimation of small-scale ABM

Gilli and Winker (2001) have been the pioneers in this field of research by applying the indirect inference method to ACE models.

Main features

- Simple ABM of the foreign exchange market (the Kirman, 1991, *ant model*);
- Dataset on the daily log-returns of the DM/US$ exchange rate;
- Two key data properties to be matched:
  - the excess kurtosis at daily frequency k_d;
  - the volatility clustering parameter α_1.

The description of the model

Two types of agents:

- fundamentalist: $E^f[\Delta S_{t+1}] = \nu(\bar{S} - S_t)$;
- chartist: $E^c[\Delta S_{t+1}] = S_t - S_{t-1}$.

An agent might change rule due to:

- random mutation, with a fixed probability ε;
- conviction, with a fixed probability δ and after having interacted with a second individual.

The market price is determined as:

$$E^m[\Delta S_{t+1}] = w_t E^f[\Delta S_{t+1}] + (1 - w_t) E^c[\Delta S_{t+1}]. \tag{5}$$

Simulated Minimum Distance

The indirect inference procedure for the estimation of the two parameters (ε, δ) solves the following minimization problem:

$$f = |\bar{\hat{k}}_d - k_d| + |\bar{\hat{\alpha}}_1 - \alpha_1| \tag{6}$$

where $\bar{\hat{k}}$ and $\bar{\hat{\alpha}}_1$ are the two data properties, but estimated on the simulated data and averaged over $10K$ Monte Carlo replications.

The algorithm

Algorithm 1 Indirect estimation procedure

Give $x^{(0)} \in \mathbf{R}^n$ starting values of parameters to be estimated
while not converged **do**
  Determine successive vectors x (defined by optimization algorithm)
  for each x **do**
    Initialize random variable generators with fixed seed
    for $i = 1 : nr\ ep$**do**
      Generate random sequences for price simulation
      Simulate price path $p^{(i)}$ and returns $r^{(i)}$
      Compute $\hat{\alpha}_{1_i}$ and k_{d_i}
    end for
    Truncate tails (10%) of the distribution of $\hat{\alpha}_{1_i}$ and k_{d_i}, $i = 1, \ldots, nr\ ep$
    Compute means $\overline{\hat{\alpha}}_1$ and $\overline{k}_d$ of truncated distributions
    Evaluate objective function
    $f = \mid \overline{k}_d - k_{d_{\mathrm{emp}}} \mid + \mid \overline{\hat{\alpha}}_1 - \alpha_{1_{\mathrm{emp}}} \mid$
  end for
end while

Figure 1: The Gilli and Winker (2001) algorithm for indirect inference of the two parameters (ε, δ).

Surface plot

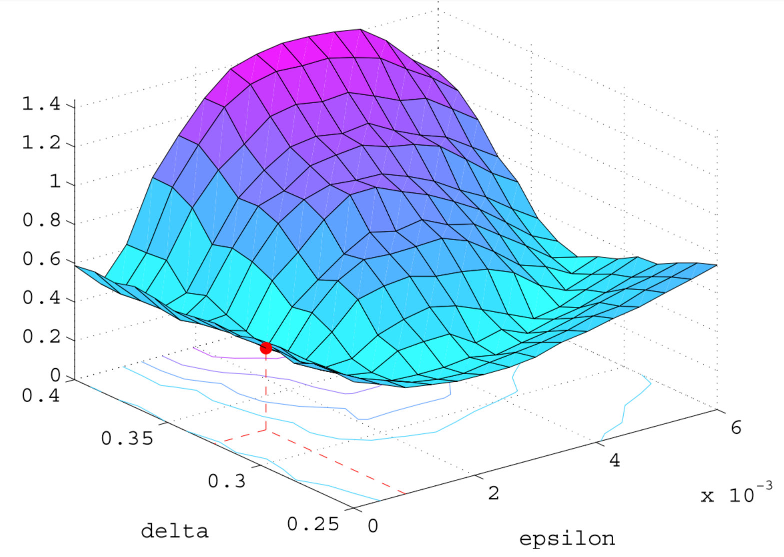


Figure 2: The Gilli and Winker (2001) surface plot of f over the (ε, δ) space.

What does this exercise tell us?

Optimal estimate is $(\varepsilon^*, \delta^*) = (0.00086, 0.325)$.

Economic implications

In the Kirman (1991) model the dynamic converges to a coexistence of fundamentalist and chartist agents if $\varepsilon < (1-\delta)(N-1)$.[3]

Hence, bringing the data to the model might suggest that the DGP of the empirical observation is better approximated by a model where the switching between the fundamentalist and chartist trading strategies matters.

The estimation seems to conclude that the relative fractions of fundamentalist and chartist traders aren't constant over time.

[3] Where $N = 100$ is the number of traders.

Limitations

Gilli and Winker (2001) recognize that:

- the two moments do not provide a sufficient description of the exchange rate dynamics (hence this estimation in hardly a test on the theory of exchange rate determination);
- further features (additional moments) shall be taken into account to better discriminate between different models;
- including other moments would however require more computational power (which was limited in 2001, but is less of a problem today for simple small-scale models).

In addition:

- the truncation of the 10% tails can be debated;
- the distance f is additive with same weights of the two objective. One can think of many possible functional forms or at different weights combinations;
- there are other parameters which have not been estimated. What about, for example, the effects of the parameter ν?

The heritage of Gilli and Winker (2001)

Gilli and Winker (2001) set the ground for the development of the literature in the estimation of ACE models.

The main algorithmic procedure adopted has been extensively used with some variations and improvements:

- Alfarano et al. (2005): simpler model with an analytical closed form solution for the distribution of returns and, therefore, direct and parametric (rather than indirect) estimation of the parameters;
- Franke and Westerhoff (2012): two models (with at most 8 parameters) estimated with respect to 9 selected moments by minimization of a weighted quadratic loss function; introduction of the *moment coverage ratio (MCR)* concept;
- Kukacka and Barunik (2017): estimation of the Brock and Hommes (1998) model by means of non-parametric simulated maximum likelihood estimation (NPSMLE).

Grazzini et al. (2017)

Introduction of Bayesian techniques to estimate ABM with applications to the Cliff and Bruten (1997) model as well as to a New Keynesian model (mid-scale).

Advantages

- does not require pre-selected moments to evaluate the model quality;
- incorporates prior knowledge and copes better with uncertainty;
- uses all the information coming from the data.

Drawbacks

- can mask identification issues by adding curvature to possibly flat likelihood functions;
- it is computationally demanding since the likelihood estimation might require as much time as the simulation of the model;
- when approximating the likelihood, the extent of the first advantage is reduced.

Bayes theorem

Bayes theorem applied to the real data, as possibly derived by the model as a DGP writes:

$$p(\theta|y) \propto \mathcal{L}(\theta; y) \cdot p(\theta) \tag{7}$$

where:

- $p(\theta|y)$ is the *posterior* distribution of the parameters θ;
- $\mathcal{L}(\theta; y) = p(y|\theta)$ is the likelihood of observing the data Y^R conditional on the simulation parameter vector θ;
- $p(\theta)$ is the *prior* distribution of the parameters θ;

Major difference with SMD

The outcome of the estimation is a multivariate distribution (univariate over each of the components of the parameter vector θ) and not a single point estimate.

Estimating the likelihood

The central issue of the Bayesian approach is the estimation of the likelihood function $\mathcal{L}(\theta; y)$.

Non-parametric estimation (slower)

$$\mathcal{L}(\theta; y) = \prod_{t=1}^{T} g(y_t^R|\theta)$$

and the estimate $\hat{g}$ is performed trough Kernel density estimation (KDE) using the simulated data after the statistical equilibrium has been reached.[4]

Parametric estimation (faster)

$$y_t = g^*(\theta) + \varepsilon_t$$

The functional form of g^* is here assumed to have some convenient form (e.g. Gaussian) but if the assumption is wrong, estimation biases are likely to arise.

[4] Ergodicity and stationarity of the model are required.

Iteration over θ

Once the likelihood is estimated, the application of Bayes theorem in equation 7 allows one to obtain a density for the posterior at one given value of θ.

But to recover the whole posterior distribution one shall sample for different $\theta \in \Theta$ (as in the standard indirect inference approach) and simulate the models at many different values.

Thus, if θ has length K and each element θ_k is sampled at J different values, we need to estimate the likelihood a number of times equal to $N_{\mathcal{L}} = J^K$.

All in all, the model requires $N = N_{\mathcal{L}} \cdot M$ independent simulations.

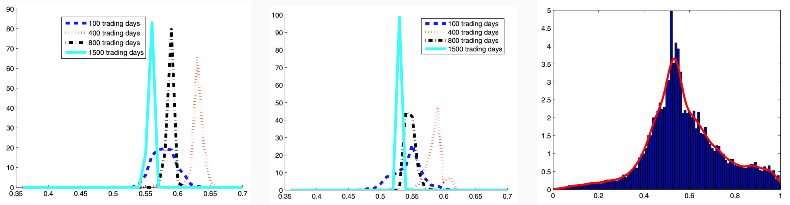


Figure 3: Posterior distribution of the parameter β under grid search and Gaussian parametric estimation, grid search with KDE and MCMC search with KDE.

Results

- Parametric Gaussian: highest precision;
- KDE with grid search: quite precise but biased;
- KDE with MCMC search: quite imprecise.

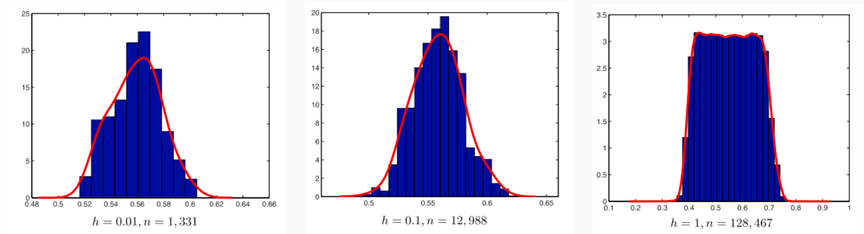


Figure 4: Posterior distribution of the parameter β under Approximate Bayesian Computation algorithm, with KDE estimated at different bin-width values.

Result

- Approximate Bayesian Computation as good as the parametric one in the best case;
- but the results are sensitive to the KDE bin-width choice;
- in particular there is a trade-off between bin-width and estimation uncertainty.

Estimation of large-scale ABM

Estimation via method of simulated moments

To our knowledge, only two mid-scale models estimated via MSM:

- Bianchi et al. (2007): macroeconomic model;
- Guerini et al. (2020): industry dynamics model.

The biggest issue remains the computational burden, which is accompanied by the *curse of dimensionality* (De Marchi, 2005).

Example: Calibrating the Dosi et al. (2015) model

- number of parameters: $K = 31$;
- number of Monte Carlo: $M = 100$;
- grid search of equispaced points for each parameter: $J = 20$;
- simulation time of a run: $T = 60s$.

Required independent simulations: $M \cdot J^K = 2.15 \cdot 10^{42}$ runs;
Required computational time: $T \cdot M \cdot J^K = 3.58 \cdot 10^{40}$ hours.

Meta-models

At this stage, is clear that the two most computationally intensive tasks for estimating an ABM are:

- running the model;
- computing the likelihood / evaluating the distance.

One approach to fasten the times of evaluation is that of replacing the model with a *surrogate* (a.k.a. meta-model) and evaluating its quality.

What is a meta-model?
Is a model/function which approximates the relation between the inputs of the complete ABM with its output. Allows one to evaluate the variation in the model outcomes over the parameter space without simulating the model.

Two possible approaches:

- parametric (see Salle and Yıldızoğlu, 2014; Dosi et al., 2018)
- non-parametric (see Lamperti et al., 2018)

Procedure

To carry out the exercise and evaluate the performance of the model the following procedure is followed:

1. draw a relatively large set of points using any standard sampling routine (e.g. NHLH) to proxy the whole parameter space;
2. employ a small random subset of these points, simulate the ABM and classify them (positive or negative label) to learn a surrogate model (i.e. a classifier system);
3. learn the surrogate model and use it to predict the probability of all the other non-classified points to be positively labelled;
4. select a small random subset of the points classified by the surrogate (with high probabilities of being positive) and evaluate them against the ABM output;
5. repeat steps 2 to 4 until a satisfactory level of performance is achieved or until a predefined number of iterations (a.k.a. *budget*) is completed.

Crucial Steps

Definition of a positive calibration

A positive calibration is a parameter vector $\theta \in \Theta$ such that the model's outcome satisfies a predefined calibration criterion.

Selection of the learning algorithm

Usage of XGBoost (a gradient boosting algorithm) that produces a statistical model that classifies/labels the sampled points.

Surrogate performance

Binary case: $\underset{\mathcal{F}}{argmax}\, F_1 = \frac{2\cdot\text{true positives}(\theta;\mathcal{F})}{2\cdot\text{true positives}(\theta;\mathcal{F})+\text{false positives}(\theta;\mathcal{F})+\text{false negatives}(\theta;\mathcal{F})}$

Real-valued case: $\underset{\mathcal{F}}{argmin}\, MSE = \frac{\sum_{i=1}^{N}(\hat{y}_i(\theta;\mathcal{F})-y_i(\theta))^2}{N}$

Out-of-sample evaluation

Use of the true positive rate $TPR = \frac{\text{number of correctly predicted positives}}{\text{number of positives in the pool}}$

Application to the Fagiolo and Dosi (2003) *island model*

Surrogate model selection criteria

- average GDP growth above 2%;
- output growth rates with AEP distribution with $b \leq 1$.
- no exercise with real data; the aim of the paper is that of evaluating the goodness of fit of the surrogate (not of the Fagiolo and Dosi (2003) *island model*).

Results

- surrogate modelling is 3750 times faster than the fully-fledge *island model*;
- with a sufficient *budget*, the surrogate model can out-of-sample correctly classify 90% of the points.

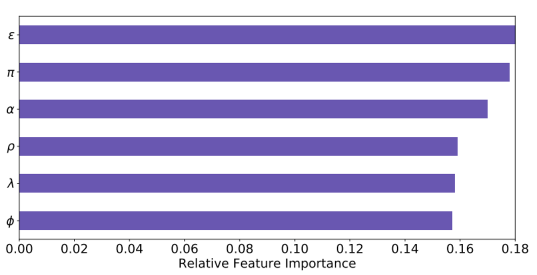


Figure 5: Importance of each parameter in shaping behaviour of the Islands model according to the specified criteria.

Results

- The meta-model can be used to decompose the variance of the explained objective as a function of to the parameters;
- here all the parameters are almost equally important;
- the parameter ε is slightly more relevant, meaning that small modifications of it can affect the calibration exercise to a larger extent.

Assessment of the strategy

Advantages

- fast due to the fact that one doesn't need to simulate the ABM many times;
- when the response surface to a variation in the parameters is very rugged, the surrogate is able to capture them;
- allows one also to assess the parameters importance and to avoid modifying parameters that do not affect the model's outcome.

Drawbacks

- the meta-model might not be a sufficient good representation of the ABM;
- when the response surface to a variation in the parameters is smooth, parametric alternatives can be more efficient;
- the surrogate is difficult to interpret: for policy analysis one needs to resort to simulation of the ABM.

New Distances and Objective Functions

New metrics for validation of ACE model

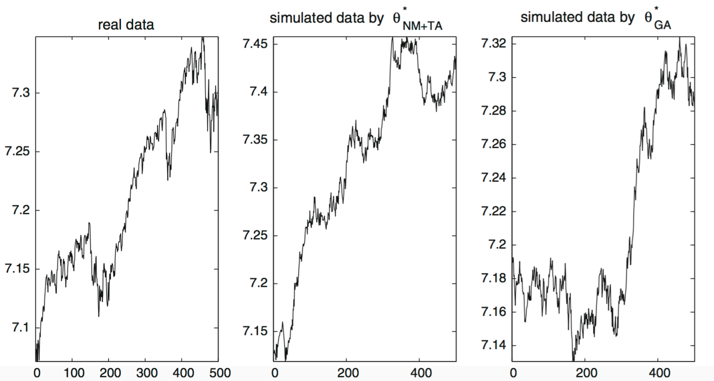
In the recent years a series of new distances to evaluate the performance of an ACE model against real data have been developed.

Among them:

- Marks (2013);
- Barde (2017, 2020)
- Lamperti (2018a)
- Guerini and Moneta (2017)

Current research is in the attempt of using these distances as minimization criteria for medium- and large-scale models in order to estimate them over a large number of parameters.

Are real and simulated data similar?



Source: Fabretti (2013)

The KL divergence

- Estimate the conditional distribution of patterns emerging from real and simulated data
- Compare the two distributions

A well known measure of the distance between distributions, with well known theoretical properties, is the Kullback-Leibler divergence:

$$D_{KL}(\mathbf{p}||\mathbf{q}) = \sum_{s \in S} p(s) \log \left(\frac{p(s)}{q(s)} \right), \tag{8}$$

The KL divergence measures the information loss when assuming that a distribution is q while the true one is p.

A **set of criteria** has been developed to capture the distance between real and simulated data:

- the State Similarity Measure (SSM – Marks, 2013);
- the Markov Information Criterion (MIC – Barde, 2017, 2020);
- the Generalized Subtracted L-divergence (GSL-div – Lamperti, 2018a,b).

Similarities and trade-offs:

- they are all built upon information theory concepts;
- But they capture slightly different features;
- And they have different computational burdens ($\text{time}_{\text{MIC}} > \text{time}_{\text{GSL}} > \text{time}_{\text{SSM}}$)

The GSL-div by Lamperti (2018a)

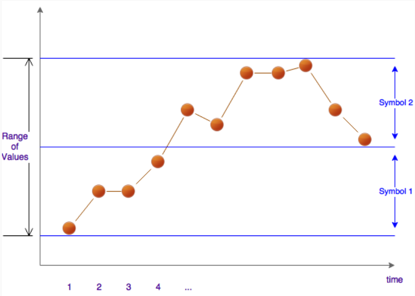
It measures the **distance between the regularity of behaviours** observed in the empirical data and that of simulated data.

It is constructed through a **simple algorithm**

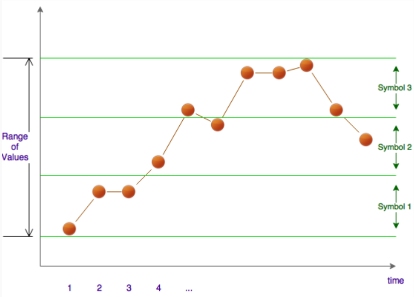
1. Time series are symbolized into b symbols
2. Sequences of symbols (words) are observed using multiple rolling window of maximal length L
3. Distributions are compared using a well known information theoretic divergence measure (L-div)
4. They are iteratively aggregated using weights that penalize short windows

Parameters b and L need to be fine-tuned.

Symbolization



(a) b (precision of symbolization) $= 2$



(b) b (precision of symbolization) $= 3$

Matching Causality by Guerini and Moneta (2017)

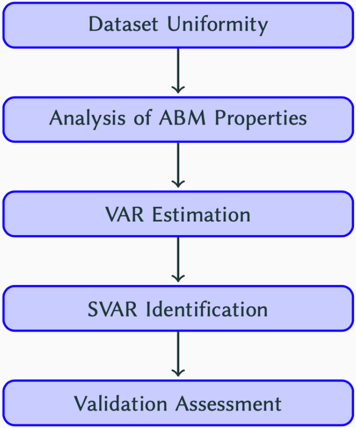
Is the mere replication of stylized facts or distribution a hard test for evaluating the model quality? If the answer is no, then one shall focus on **causal relationships**.

The intuition is that models with different causal structures, might generate equivalent statistical properties (i.e. the same stylized facts).

How to discriminate between them?

- estimate the causal structures emerging from real-word data and model-generated data;
- compare them in a meaningful way.

The procedure



Ergodicity and stationarity can be tested by batteries of Kolmogorov-Smirnov tests.

The SVAR identification can be done by modern causal search algorithms with relatively mild assumptions, without imposing **a-priori restriction**.

The **similarity** is computed using signs and magnitudes of the causal relationships estimated on the real and on the simulated data.

Application has been done at the moment on the Dosi et al. (2015) model.

Currently on-going: a comparison of Grazzini et al. (2017) and Guerini and Moneta (2017) on the Delli Gatti and Desiderio (2015) model.

Conclusions

Overall Evaluation

After almost 20 years from the Gilli and Winker (2001) seminal contribution:

- computational power has hugely increased expanding the possibilities for estimation and calibration of ACE models;
- many alternative nuances of the estimation method are now available, but their applications by ACE modellers is still scant.

The quest for the *optimal estimation strategy*?
Probably not, as all methods have advantages and drawbacks.

Future research will likely aim at:

- specifying the statistical properties of the different estimators as well as the conditions for their usage to better define the applicability domains of the different methods;
- writing a broad open-source software embedding most of the methodologies to allow ABM economists and policy-makers to easily implement them.

A Validation Cookbook for AB Modellers

Ideal procedure for *descriptive models*:

1. **input validation**: build the model by grounding behavioural rules and interaction mechanisms on the empirical or experimental literature
2. **estimation**: decide the real-world data, the moments to match and estimate the model with the preferred method
3. **estimation robustness**: run Monte Carlo simulations and check if the model dynamics resembles the real world one according to some distance measure
4. **stylized facts replication**: check which stylized facts the model can replicate (in addition to the ones used for the estimation)
5. **causal structure validation**: check if the model causal structure resembles the one that emerges from the real data (if the model has policy implications, this is particularly relevant)
6. **sensitivity analysis**: run local sensitivity analysis or use a surrogate to globally explore the parameter space
7. **policy analysis**: run policy exercises and compare results to the baseline parametrization

Thanks for the attention.

References 1

Alfarano, S., T. Lux, and F. Wagner (2005). Estimation of agent-based models: The case of an asymmetric herding model. *Computational Economics 26*(1), 19–49.

Barde, S. (2017, Aug). A practical, accurate, information criterion for nth order markov processes. *Computational Economics 50*(2), 281–324.

Barde, S. (2020). Macroeconomic simulation comparison with a multivariate extension of the markov information criterion. *Journal of Economic Dynamics and Control 111*, 103795.

Bianchi, C., P. Cirillo, M. Gallegati, and P. A. Vagliasindi (2007). Validating and calibrating agent-based models: A case study. *Computational Economics 30*(3), 245–264.

Brock, W. (1999). Scaling in economics: a reader's guide. *Industrial and Corporate Change 8*(3), 409–446.

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

Buchanan, M. (2012). It's a (stylized) fact! *Nature Physics 8*(1), 3.

Cliff, D. and J. Bruten (1997). Minimal-intelligence agents for bargaining behaviors in market-based environments. Technical report.

De Marchi, S. (2005). *Computational and mathematical modeling in the social sciences.* Cambridge University Press.

Delli Gatti, D. and S. Desiderio (2015). Monetary policy experiments in an agent-based model with financial frictions. *Journal of Economic Interaction and Coordination 10*(2), 265–286.

Dosi, G., G. Fagiolo, M. Napoletano, A. Roventini, and T. Treibich (2015). Fiscal and monetary policies in complex evolving economies. *Journal of Economic Dynamics and Control 52*, 166–189.

Dosi, G., M. C. Pereira, and M. E. Virgillito (2018). On the robustness of the fat-tailed distribution of firm growth rates: a global sensitivity analysis. *Journal of Economic Interaction and Coordination 13*(1), 173–193.

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.

Fagiolo, G. and G. Dosi (2003). Exploitation, exploration and innovation in a model of endogenous growth with locally interacting agents. *Structural Change and Economic Dynamics 14*(3), 237–273.

Fagiolo, G., M. Guerini, F. Lamperti, A. Moneta, and A. Roventini (2019). Validation of Agent-Based Models in Economics and Finance. pp. 763–787.

Fagiolo, G. and A. Roventini (2017). Macroeconomic policy in DSGE and agent-based models redux: New developments and challenges ahead. *Jasss 20*(1).

Franke, R. and F. Westerhoff (2012). Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. *Journal of Economic Dynamics and Control 36*(8), 1193–1211.

Gilli, M. and P. Winker (2001). Indirect Estimation of the Parameters of Agent Based Models of Financial Markets. *SSRN Electronic Journal* (November).

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect Inference C. Gourieroux; A. Monfort; E. Renault Journal of Applied Econometrics , Vol. 8, Supplement: Special Issue on Econometric Inference Using Simulation Techniques. (Dec., 1993), pp. S85-S118. *Journal of Applied Econometrics 8.*

Grazzini, J., M. G. Richiardi, and M. Tsionas (2017). Bayesian estimation of agent-based models. *Journal of Economic Dynamics and Control 77*, 26–47.

Guerini, M., P. Harting, and M. Napoletano (2020). Governance structure, technical change and industry competition. Technical report.

Guerini, M. and A. Moneta (2017). A method for agent-based models validation. *Journal of Economic Dynamics and Control 82.*

Hsiao, C. (1989). Consistent estimation for some nonlinear errors-in-variables models. *Journal of Econometrics 41*(1), 159–185.

Kaldor, N. (1957, 12). A Model of Economic Growth. *The Economic Journal 67*(268), 591–624.

Kirman, A. (1991). Epidemics of opinion and speculative bubbles in financial matkets. *Money and financial markets chap 17.*

Kukacka, J. and J. Barunik (2017). Estimation of financial agent-based models with simulated maximum likelihood. *Journal of Economic Dynamics and Control 85*, 21–45.

Lamperti, F. (2018a). An information theoretic criterion for empirical validation of simulation models. *Econometrics and Statistics 5*, 83–106.

Lamperti, F. (2018b). Empirical validation of simulated models through the gsl-div: an illustrative application. *Journal of Economic Interaction and Coordination 13*(1), 143–171.

Lamperti, F., A. Roventini, and A. Sani (2018). Agent-based model calibration using machine learning surrogates. *Journal of Economic Dynamics and Control 90*, 366–389.

Marks, R. E. (2013). Validation and model selection: Three similarity measures compared. *Complexity Economics 2*(1), 41–61.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica 57*(5), 995.

Naylor, T. H. and J. M. Finger (1967). Verification of Computer Simulation Models. *Management Science 14*(2), B–92–B–101.

Salle, I. and M. Yıldızoğlu (2014). Efficient sampling and meta-modeling for computational economic models. *Computational Economics 44*(4), 507–536.

Yilmaz, L. (2006). Validation and verification of social processes within agent-based computational organization models. *Computational & Mathematical Organization Theory 12*(4), 283–312.