Validating and Selecting Agent-Based Models

C^3 International Biannual Symposium
26 November 2015
University of Sydney

Robert E. Marks
Economics, UNSW Australia

robert.marks@gmail.com
www.agsm.edu.au/bobm

OUTLINE

1. Introduction
2. Sufficiency and Necessity
3. Validation
4. Validation as Model Selection
    - The State Similarity Measure (SSM)

1. Introduction

Computer scientists are concerned with *finding solutions* to issues such as market design, whereas —

Social scientists in general and economists in particular have been concerned with *explaining and predicting* social phenomena. (Marks 2012)

(This is also true of other scientists, such as ALIFE researchers (Bentley 2013) and the 2013 Chemistry Nobel laureates, who use simulations to model real-world phenomena.)

Both of these approaches demand sufficiency, but scientists (or at any rate economists) also want necessity (its importance is moot):

Not just: "This is a solution"

but also: "This is the set of all possible solutions."

Traditional Economic Methods

A certain logic:

- observe a real-world phenomenon
- identify a need to explain and understand it
- build a methematical, closed-form model, with simplifying assumptions to allow its solution
- manipulate the model to obtain sufficient and necessary conditions for the observed phenomenon
- perhaps relax a simplifying assumption or two and ask how the model changes

This has focussed on equilibria or steady-states, precluding study of out-of-equiibrium or dynamic phenomena.

Simulation can overcome these restrictions, but at a cost.

Judd's ideas (2006)

"Far better an approximate answer to the right question ... than an exact answer to the wrong question."

— John Tukey, 1962.

That is, economists face a tradeoff between:

the numerical errors of computational work
and
the specification errors of analytically tractable models.

And perhaps also between: sufficiency and necessity.

2. Sufficiency and Necessity

Simulations demonstrate: *existence* and *sufficiency,*

but only in certain circumstances *necessity.*

Simulations can demonstrate the untruth of a proposition,

but not provide proofs or theorems,

simulations cannot provide generality.

What, never?

Does this matter?

Formal Simulation

Mathematical "model A" comprises the conjunction $(a_1 \wedge a_2 \wedge a_3 \cdots \wedge a_n)$, where $\wedge$ means "AND", and the a_i denote the elements (equations, parameters, initial conditions, etc) that constitute the model.

Sufficiency: If model A exhibits the desired target behaviour B, then model A is sufficient to obtain exhibited behaviour B:
$A \Rightarrow B$

Thus, any model that exhibits the desired behaviour is sufficient, and demonstrates one conjunction of conditions (or *model*, or solution) under which the behaviour can be simulated.

But if there are several such models, how can we choose among them? And what is the necessary set $\mathcal{N}$ of all such conjunctions (models)?

Necessity

Necessity: **Only those models A belonging to the set of necessary models $\mathcal{N}$ exhibit target behaviour B.**

That is, $(A \in \mathcal{N}) \Rightarrow B$, and $(D \notin \mathcal{N}) \not\Rightarrow B$.

A hard challenge: to determine the set of necessary models, $\mathcal{N}$.

Since each model is not simple: $A = (a_1 \wedge a_2 \wedge a_3 \cdots \wedge a_n)$, searching for the set $\mathcal{N}$ of necessary models means searching in a high-dimensional space, with no guarantee of continuity, and a possible large number of non-linear interactions among elements.

Explanation **could be aided from examination of what the models in the necessary set $\mathcal{N}$ have in common.**

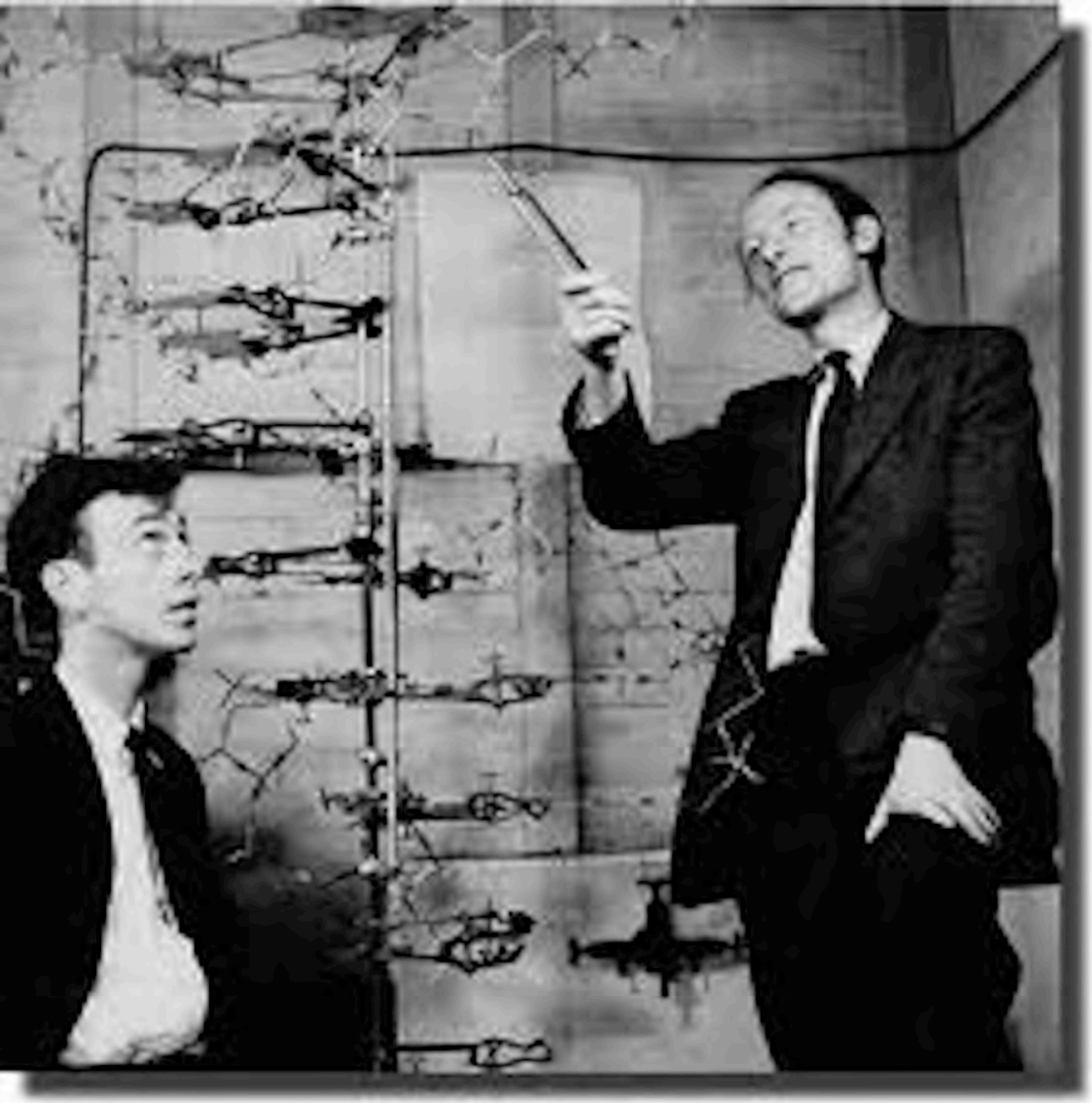
Lack of Necessity Means ...

For instance, if $D \not\Rightarrow B$, it does not mean that *all* elements a_i of model D are invalid or wrong, only their conjunction, that is, model D.

It might be only a single element a_k that precludes model D from exhibiting behaviour B.

But determining whether this is so and which is the offending element a_k is a costly exercise, in general, for the simulator.

Without clear knowledge of the boundaries of the set $\mathcal{N}$ of necessary models, it is difficult to generalise from simulations.



Simulation Can Sometimes Demonstrate Necessity . . .

Only when the set $\mathcal{N}$ of necessary models is known to be small (such as in the case of DNA structure in 1953 when Watson & Crick were searching for it) is it relatively easy to use simulation to derive necessity.

Watson & Crick had much information about the properties of DNA (from others):

when they hit on the simulation we know as the "double helix", they knew it was right.

But still "*A* structure ...", not "*The* structure" in the title of their 1953 *Nature* paper.

(And Kepler's 1605 ellipses v. Ptolomy's epicycles?)

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

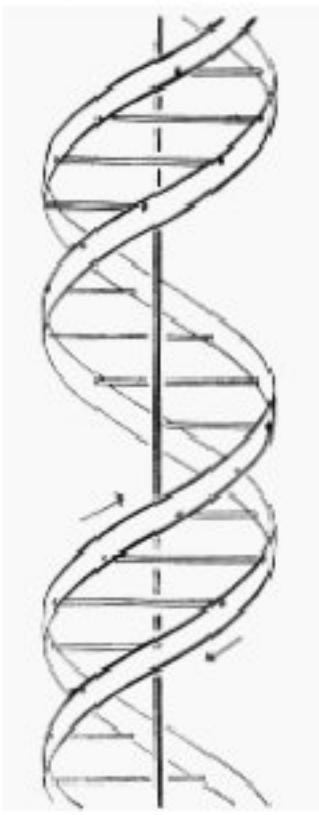
We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three inter-twined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion. this structure is unsatisfactory for two reasons:

(1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions. namely, that each chain consists of phosphate diester groups joining ß-D-deoxy-ribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow righthanded helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the cbains together. The vertical line marks the fibre axis

Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3-4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain, does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain, is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in time following, communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereo-chemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. *M.* H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J.D.W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J.D. WATSON
F.H. C. CRICK

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge. April 2.

[1] Pauling, L., and Corey, R. B. nature, 171, 346 (1953); Proc, U.S. Nat Acad. Sci., 39, 84 (1953).

[2] Furberg, S., Acta Chem. Scand., 6, 634 (1952).

[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., Biochim, et Biophys. Acta, 9402 (1952).

[4] Wyatt, G.R. J. Gen. Physiol., 36 201 (1952).

[5] Astbury, W.T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947)

[6] Wilkins, M. H. F. and Randall, J. T. Biochim. et. Biophys. Acta, 10, 102 (1953).

Verification & Validation — "Assurance"

Verification (or internal validity): is the simulation working as you want it to:

— is it "doing the thing right?"

Validation: is the model used in the simulation correct?

— is it "doing the right thing?"

To Verify: use a suite of tests, and run them every time you change the simulation code — to verify the changes have not introduced extra bugs.

See: D.F. Midgley, Marks R.E., and Kunchamwar D. (2007)

3. Validation

For whom?

With regard to what?

A good simulation is one that achieves its goals:

- **to explore**
- **to predict**
- **to explain**

Or

- **what is?**
- **what could be?**
- **what should be?**

Validation

To the extent that the social sciences are concerned with real-world, historical phenomena,

any simulations must be verified (no bugs) and validated (does the model provide behaviour which matches the stylised facts of the historical phenomenon?)

Midgley et al: verification + validation = *assurance*

Back-predictions.

Docking.

Agent-Based Models → Generative Explanation:

Generative explanation (Epstein 2006):

"If you haven't grown it, you haven't explained its *emergence*."

(But does growth ⇒ explanation?)

To answer: how could the autonomous, local interactions of heterogeneous boundedly rational agents generate the observed regularity (that emerges)?

— Generative sufficiency is a necessary but not sufficient condition for explanation. Each realisation is a strict deduction.

Grüne-Yanoff (2006) argues to distinguish *functional explanations* (easier for simulators) from *causal explanations* (much less achievable for social scientists).

Prediction

Prediction only requires sufficiency, not necessity.

Two recent examples of using simulation models to predict:

1. **Katie Bentley and associates built an AB model of angiogenesis, which turned out to predict behaviour *in silico* never before seen *in vivo*, but later observed (Bentley et al. 2009, 2013).**

2. **The 2013 Nobel Prize for Chemistry was awarded for work in the simulation of chemical reactions, using both classical and quantum mechanisms, which has made it possible to "optimize catalysts, drugs, and solar cells." (Royal Academy 2013).**

Correct prediction $\Rightarrow$ validity, but may be too strict.
See Troitzsch (2004) and Hassan et al. (2013).

Validation

Moss & Edmonds (2005): for AB models at least two stages of empirical validation.

1. **the micro-validation of the behaviour of the individual agents in the model, by reference to data on individual behaviour.**
2. **macrovalidation of the model's aggregate or emergent behaviour when individual agents interact, by reference to aggregate time series.**

   with the emergence of novel behaviour, possible surprise and possible highly non-standard behaviour, difficult to verify using standard statistical methods.

   $\therefore$ only qualitative validation judgments might be possible.

Formalisation of Validation , Outputs Graphically ...

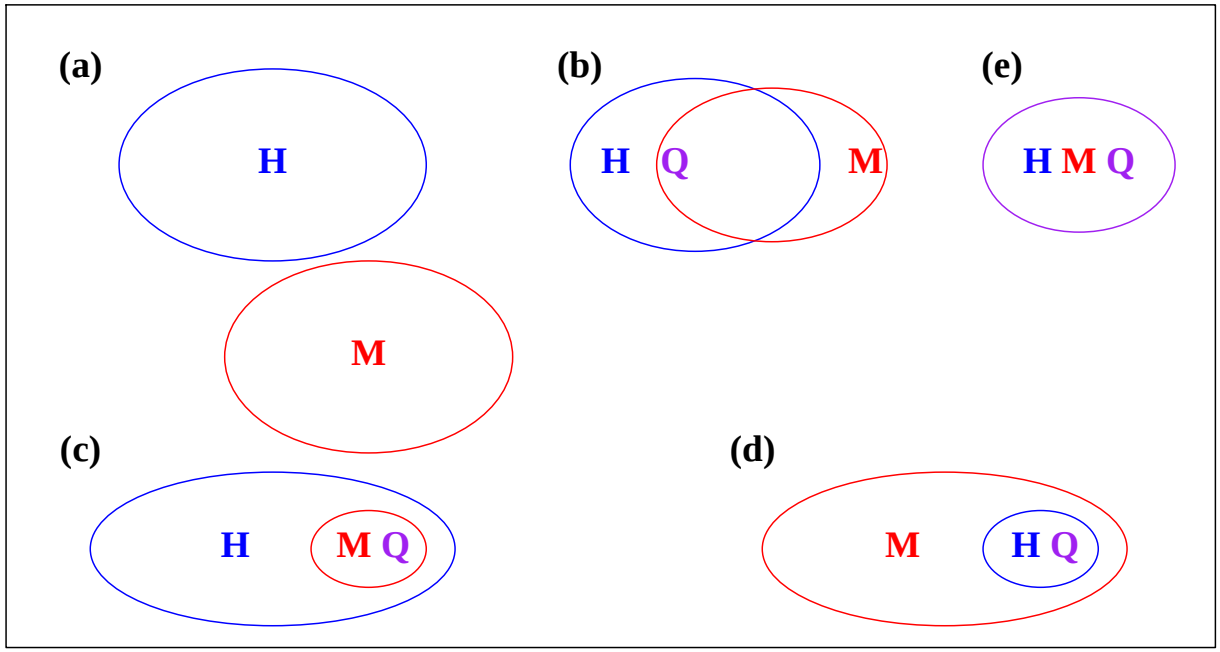


Figure 2: Validity relationships (from Marks 2007).

a. **useless**
b. **useful, but incomplete and inaccurate**
c. **accurate but incomplete**
d. **complete but inaccurate ← possibly the best to aim for**
e. **complete and accurate**

Modelling Goals

One goal: to construct and calibrate the model so that Case e:

$M \approx Q \approx H$: there are very few historically observed behaviours that the model does not exhibit,

and there are very few exhibited behaviours that do not occur historically.

The model is close to being both complete and accurate.

In practice, a modeller might be happier to achieve Case d., where the model is complete (and hence provides sufficiency for all observed historical phenomena), but not accurate.

Not least to accommodate later real-world observations.

Fagiolo et al. (2007) on Validation of AB Models

AB models can be characterized as:

- bottom-up models (unlike e.g. Systems Dynamics simulation models, or closed-form models)
- heterogeneous agents (endowments, properties, memory, rationality, etc.)
- boundedly rational, usually with adaptive expectations
- networked direct interactions.

Closer to dynamic, decentralized markets and economies than traditional models.

Fagiolo 2

And yet marked reluctance to use AB models. Why? Four key problems:

1. no common set of the heterogeneous AB models previously developed
2. (hence) lack of comparability across these models with high degrees of freedom, hence a wide range of outputs, together with lack of necessity.
3. lack of standard techniques for constructing and analyzing AB models, but see Grimm et al. (2010),
4. the "problematic" relationship between AB models and empirical data $\leftarrow$ this is *validation*.

And see Waldherr and Wijermans (2013) for further on this.

Fagiolo 3 — Issues with Empirical Validation

Comparing historical data with generated outputs.

1. **The world is complex: a trade-off between complexity in modelling ("concretization") and reductionism ("isolation"): where to draw the line in modelling? Realism v. tractability.**
2. **Friedman (1953) argued that realism was not necessary so long as the output allowed accurate prediction ("instrumentalism"), but others seek realism in the model and its assumptions as well as accuracy.**
3. **How wedded should the modeller be to a priori assumptions (about the goals of agents, say)? Or should all aspects of the model be available ("pluralism")?**
4. **Importantly: how to choose which of several models is best (the "identification" or "under-determination" problem). $\leftarrow$ our later focus**

Fagiolo 4

Three alternative approaches:

1. the indirect calibration approach
2. the Werker-Brenner approach
3. the history-friendly approach, and

And the fourth, our

4. the State Similarity Measure (Marks 2013)

Two Kinds of ABM

We can think of two kinds of ABM:

1. *demonstrative* ABM models

   These models demonstrate principles, rather than tracking historical phenomena. A demonstrative ABM is an existence proof.
   Examples: Schelling's Segregation Game, my Boys and Girls NetLogo model, my Emergence of Risk Neutrality, and others

2. *descriptive* ABM models.

   These models attempt to derive sufficient conditions to match historical phenomena, as reflected in historical data. This requires validation (model choice). Ideally, to predict.
   Examples: Midgley et al. modelling brand rivalry, ALIFE models, etc

4. Validation as Model Selection

"All models are wrong, but some are useful" — Box (1976).

"Are the models wrong in ways that are central to the questions you want to ask, or are they wrong in ways that aren't so central?" Lars Hansen, Nobel Economics Laureate 2013.

Anderson & Burnham (2002) make a strong case for validation as *model selection*: for the researcher generating a selection of models, and choosing the model which loses the least *information* compared to reality (the Historical data).

How to select a "best approximating model" from the set? Anderson & Burnham review and use Akaike's information criterion (AIC).

Akaike's Information Criterion (AIC)

Considering AIC to be an extension of R. A. Fisher's likelihood theory, Akaike (1973) found a simple relationship between Kullback-Leibler "distance" or "information" and Fisher's maximized log-likelihood function.

$\rightarrow$ very general methodology for selecting a parsimonious approximating model.

Can think of modelling as being directed towards finding a good approximating model of the information encoded in the empirical, historical data.

Information about the process under study exists in the data. Want to express this information in a model: more compact, and understandable.

The role of a good model is to filter the historical data so as to separate information from noise.

A new technique: the State Similarity Measure

Our new technique, the State Similarity Measure (SSM), tackles the fourth core issue of Fagiolo et al. (2007): validating agent-based models using historical data. (Marks 2013).

Compared to their three methods (indirect calibration, the Werker-Brenner approach, and the history-friendly approach), SSM focusses on:

- the micro-level output
- an empirical comparison of model output v. history
- leading to a choice of model that best fits, to resolve any identification (or under-determination) issues.

The SSM can include a Monte Carlo simulation to eliminate random observations (Marks 2014).

One issue: heterogenous agents, time-series price, defining the states of the market

Two reasons to compare such model output against history:

1. To choose better parameter values, to "calibrate" or (more formally) "estimate" the model against the historical record.
2. To choose the "best" model from a selection of possible models (different structures, parameter values, etc)

We are interested in the second, having used machine learning (the GA) to calibrate the model parameters in order to improve each agent's weekly profits (instead of fitting to history) in our agent-based model.

Figure 1 shows Historical data from a U.S. supermarket chain's sales of (heterogeneous) brands of sealed, ground coffee, by week in one city (Midgley et al. 1997).

Historical Data: Prices and Volumes in Chain 1

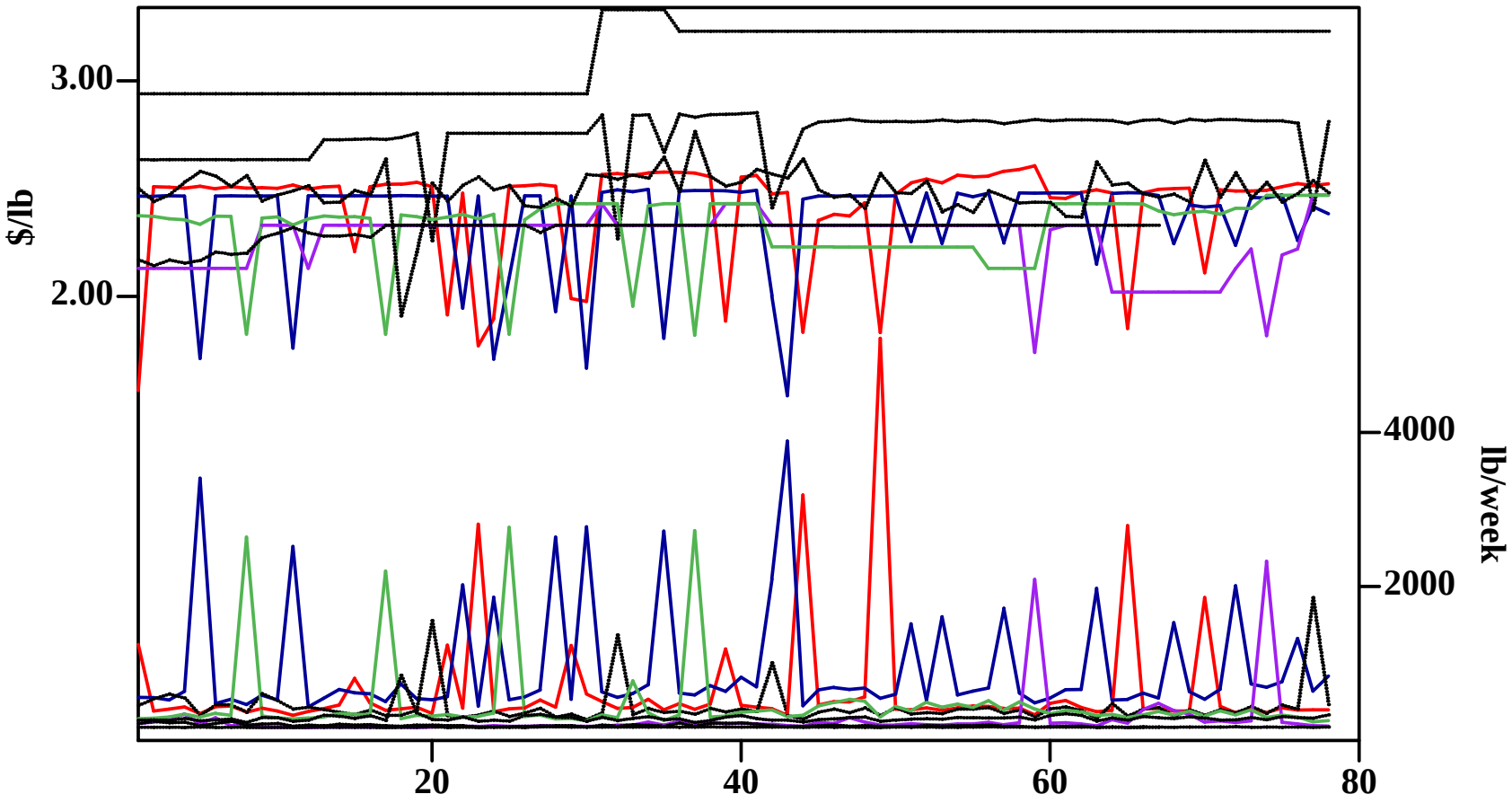


Figure 1: Weekly Sales and Prices (Source: Midgley et al. 1997)

Stylised facts of the Historical data:

1. Much movement in the interactive prices and volumes of four strategic brands, and *strategic* means?

2. For these four (coloured) brands, high prices (and low volumes) are punctuated by a low price (and a high volume).

3. Another five (non-strategic) brands exhibit stable (high) prices and (low) volumes.

In addition, the competition is not open slather: the supermarket chain imposes some restrictions on the timing and identity of the discounting brands.

A Model of Strategic Interaction

We assume that the price P_{bw} of brand b in week w is a function of the state of the market M_{w-1} at week $w-1$, where M_{w-1} in turn might be a product of the weekly prices S_{w-j} of all brands over several weeks:

$$P_{bw} = f_b(M_{w-1}) = f_b(S_{w-1} \times S_{w-2} \times S_{w-3} \cdots)$$

Earlier in the research program undertaken with Midgley et al., we used the Genetic Algorithm to search for "better" (i.e. more profitable) brand-specific mappings, f_b, from market state to pricing action.

And derived the parameters of the models, and derived their simulated behaviour, as time-series patterns (below).

Partitioning the Data

A curse of dimensionality: each brand can price anywhere between \$1.50 and \$3.40 per pound: 190 price points. Consider three stragic brands only.

The first coarsening:
Marks (1998) explores partitioning while maximising information (using an entropy measure). Finds that dichotomous partition is sufficient.

Here: use symmetric dichotomous partitioning: a brand's price is labelled 0 if above its midpoint, else 1 below.

The second coarsening:
Consider three depths of memory:
with 1-week memory, three brands, each pricing Low or High:
$2^3 = 8$ possible states;
with 2-week memory: $8^2 = 64$ possible state;
with 3-week memory: $64^2 = 512$ possible states.

Dichotomous Symmetric Price Partitioning of History

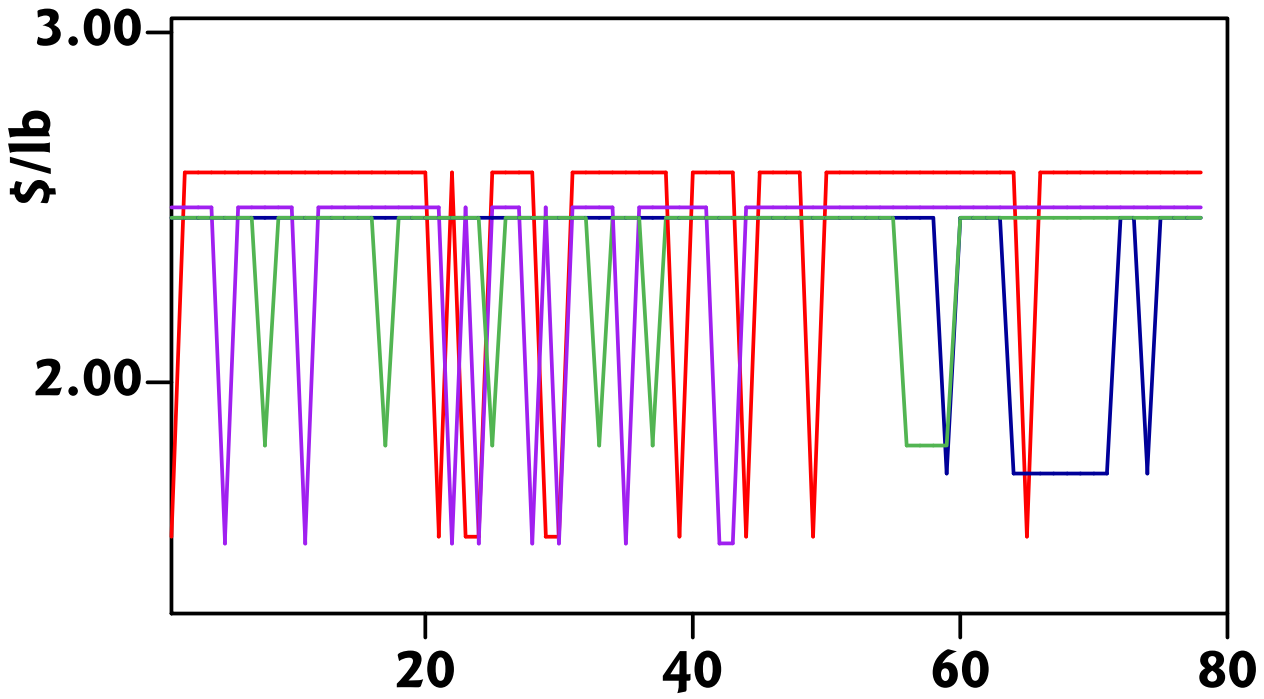


Figure 2: Partitioned Historical Prices of the Four Chain-One Brands

Example of a Simulated Oligopoly (Marks et al. 1995)

Simulating rivalry between the three asymmetric brands: 1, 2, and 5, Folgers, Maxwell House, and Chock Full O Nuts.

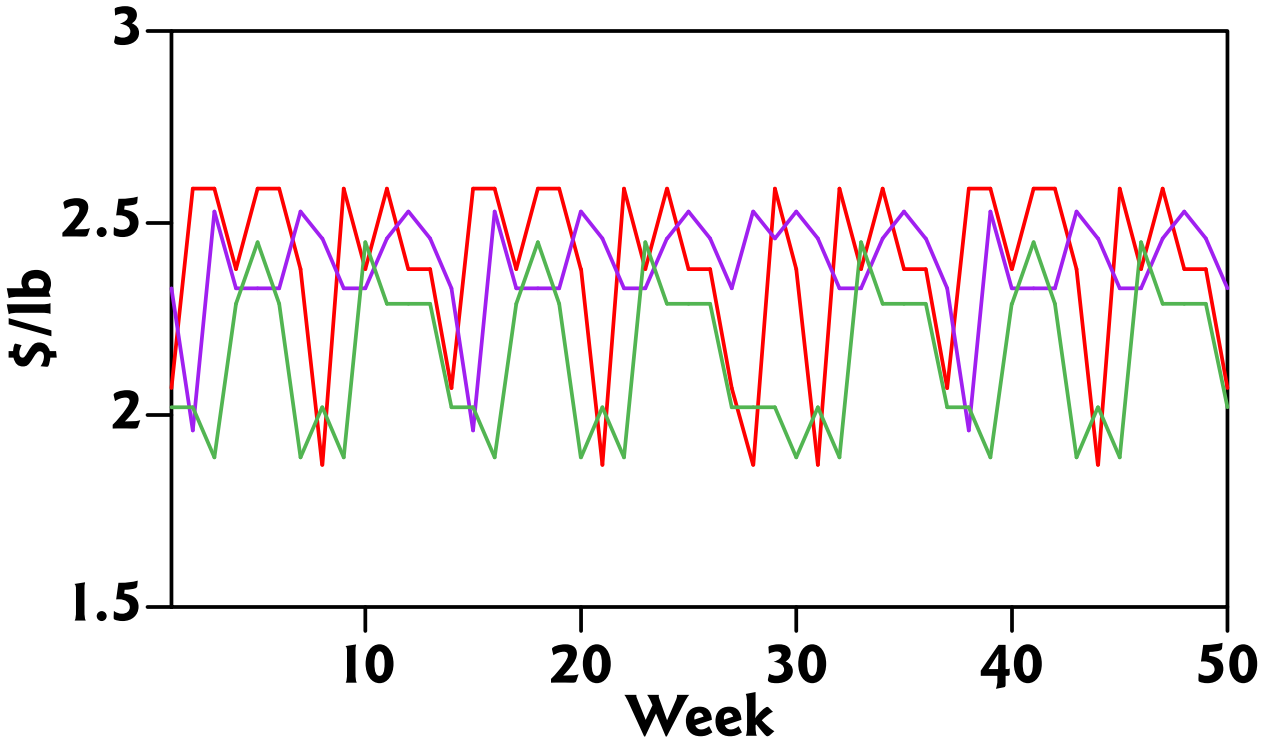


Figure 3: **Example of a Simulated Oligopoly (Marks et al. 1995)**

Three brands, one-week memory, 50 weeks observed

Table 1 shows the observed distribution of states in the Historical Chain 1, and in the three models: 11, 26a, and 26b.

State	History	Model 11	Model 26a	Model 26b
000	32	0	30	20
001	2	18	11	10
010	6	15	3	7
011	1	0	0	0
100	7	16	5	12
101	0	0	0	0
110	2	0	1	1
111	0	1	0	0
Total	50	50	50	50

So: how close are the three models to History?

Table 1: The observed frequencies of the 8 states over 50 weeks.

Results using the SSM

We derive the distance between two sets of time-series using the SSM by calculating the sum of absolute differences (city-block metric) in observed window states between the two set. (We include three brands and three-week memory.) So what?

First, the greater the sum, the more distant the two sets of time-series.

Second, we can calculate the maximum size of the summed difference: zero intersection between the two sets (no states in common) implies a measure of $2 \times S$ where S is the number of possible window states, from the data.

Third, we can derive some statistics to show that any pair of sets in not likely to include random series.

SSM Distances Between Historical Chain 1 and Three Models

	History	Model 11	Model 26a	Model 26b
History	0	92*	54	68
Model 11	92*	0	88*	80*
Model 26a	54	88*	0	60
Model 26b	68	80*	60	0

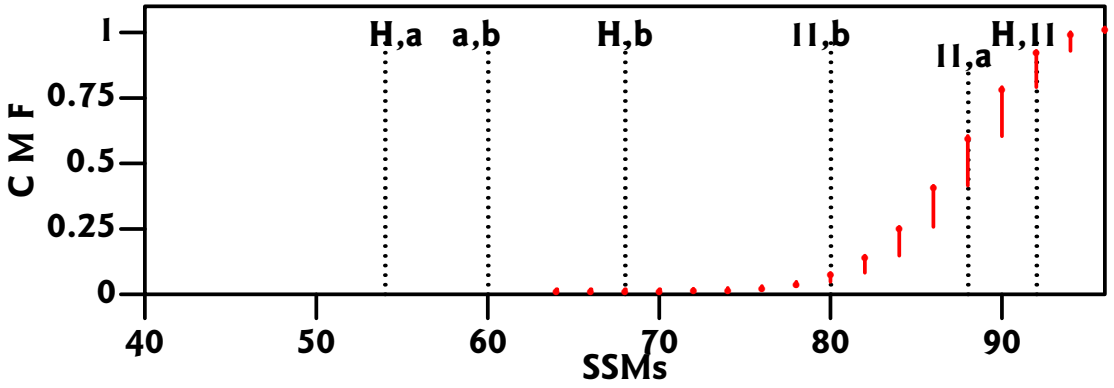
Table 5: Distances Between History and Three Models (with 3 Brands, 3-week memory)

Here, S, the maximum number of states = 48, so the maximum distance apart is 96.

We see that Model 26a is closest to Historical Chain 1, closer than it is to Model 26b; we also see that Model 11 is very distant from History, possibly into randomness.

Null Hypothesis: each of the two sets of time series is random.

Testing for Randomness



The red lines are the CMF of pairs of sets of random series (3 series, 48 observations) from 100,000 Monte Carlo parameter bootstraps.

The one-sided confidence interval at 1% corresponds to a SSM of 76, and at 5% 80.

Cannot reject the null hypothesis (random sets) for History and Model II; reject the null (random) hypothesis for all other pairs. (In the Figure, * : cannot reject the null at the 5% level.)

Conclusions — the SSM

The SSM is a true metric sufficient to allow us measure the degree of similarity between two *sets* of time-series which embody dynamic responses.

The SSM has been developed to allow us to measure the extent to which a simulation model that has been calibrated on some other criterion (e.g. weekly profitability) is similar to historical sets of time-series.

The SSM will allow us to identify which of several models' outputs is closest to history, to determine which model has captured most information of the historical time series.

Validation can be used to select the best simulation model.

References

[1] Akaike H. (1973), "Information theory as an extension of the maximum likelihood principle," in B.N. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory.* Budapest: Akademiai Kiado, pp. 267–281.

[2] R. Axelrod (2006), Advancing the Art of Simulation in the Social Sciences, in J.-P. Rennard (ed.), *Handbook of Research on Nature-Inspired Computing for Economy and Management,* (Hershey, PA: Idea Group Inc.)

[3] Katie Bentley, Holger Gerhardt and Paul A. Bates (2009), Agent-based simulation of Notch mediated tip cell selection in angiogenic sprout initialisation, *Journal of Theoretical Biology,* 250, 1 (2009) 25.

[4] Katie Bentley, Martin Jones, Bert Cruys (2013), Predicting the future: Towards symbiotic computational and experimental angiogenesis research, *Experimental Cell Research* 3 19 (2013): 1240–1246.

[5] Burnham K.P. and Anderson D.R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd. ed., New York: Springer.

[6] J.M. Epstein (2006), Remarks on the Foundations of Agent-Based Generative Social Science, *Handbook of Computational Economics, Volume 2: Agent-Based Modeling*, edited by Leigh Tesfatsion and Kenneth L. Judd, Amsterdam: Elsevier Science.

[7] Fagiolo G., Moneta A., and Windrum P. (2007), "A critical guide to empirical validation of agent-based models in economics: methodologies, procedures, and open problems," *Computational Economics,* 30(3): 195–226.

[8] Fisher R.A. (1922), "On the mathematical foundations of theoretical statistics." *Royal Society of London. Philosophical Transactions* (Series A) 222: 309–368.

[9] Friedman M. (1953), *Essays in Positive Economics*, University of Chicago Press.

[10] Grimm V., Berger U., DeAngelis D. L., Polhill J. G., Giske J., and Railsback S. F. (2010) The ODD protocol: A review and first update. *Ecological Modelling,* 221(23): 2760–2768.

[11] T. Grüne-Yanoff (2006), The Explanatory Potential of Artificial Societies, *Models and Simulations*.

[12] Samer Hassan, Javier Arroyo, José Manuel Galán, Luis Antunes and Juan Pavón (2013), Asking the Oracle: Introducing Forecasting Principles into Agent-Based Modelling, *Journal of Artificial Societies and Social Simulation* 16 (3) 13

[13] K. L. Judd (2006), Computationally Intensive Analyses in Economics, *Handbook of Computational Economics, Volume 2: Agent-Based Modeling*, ed. by Leigh Tesfatsion & Kenneth L. Judd, Amsterdam: Elsevier Science, Ch. 2.

[14] Charles M. Macal (2013), "On Validating Multi-Agent System Applications" Keynote Address, MABS 2013, 14th International Workshop on Multi-Agent-Based Simulation, May 7.

[15] R.E. Marks (1998) Evolved perception and behaviour in oligopolies, *Journal of Economic Dynamics and Control*, 22(8–9): 1209–1233, July.

[16] Marks R.E. (2007), "Validating simulation models: a general framework and four applied examples," *Computational Economics,* 30(3): 265–290, October

[17] R.E. Marks (2012), "Analysis and synthesis: multi-agent systems in the social sciences," *The Knowledge Engineering Review,* 27(2): 123–136,

[18] Marks R.E. (2013), "Validation and Model Selection: Three Similarity Measures Compared," *Complexity Economics*, Vol. 2, 01/2013.
http://www.agsm.edu.au/bobm/papers/parispaper.pdf

[19] Marks, R.E. (2014). "Monte Carlo," in *The Palgrave Encyclopædia of Strategic Management*, edited by David Teece and Mie Augier, London: Palgrave.
http://www.agsm.edu.au/bobm/papers/MonteCarloPESM-longer.pdf

[20] R.E. Marks, D.F. Midgley, and L.G. Cooper (1995) Adaptive behavior in an oligopoly, *Evolutionary Algorithms in Management Applications*, ed. by J. Biethahn and V. Nissen, (Berlin: Springer-Verlag), pp. 225–239.
http://www.agsm.edu.au/bobm/papers/marks-midgley-cooper-95.pdf

[21] D.F. Midgley, R.E. Marks, and L.G. Cooper (1997), Breeding competitive strategies, *Management Science*, 43: 257–275.

[22] D.F. Midgley, Marks R.E., and Kunchamwar D. (2007) The Building and Assurance of Agent-Based Models: An Example and Challenge to the Field, *Journal of Business Research*, Special Issue: Complexities in Markets, 60: 884–893.

[23] Moss, S. and Edmonds, B. (2005), Sociology and simulation: statistical and quantitative cross-validation, *American Journal of Sociology*, 110(4): 1095–1131.

[24] The Royal Swedish Academy of Sciences (2013), Taking experiment to cyberspace, Nobel Prize in Chemistry 2013.

[25] Schelling, T.C. (1971), Dynamic models of segregation, *Journal of Mathematical Psychology*, 1: 143–186.

[26] Klaus G. Troitzsch (2004), "Validating Simulation Models," *Proceedings of the 18th European Simulation Multiconference*, SCS Europe.

[27] John Tukey (1962), The future of data analysis, *Annals of Mathematical Statistics* 33 (1).

[28] Annie Waldherr and Nanda Wijermans (2013), Communicating Social Simulation Models to Sceptical Minds, *Journal of Artificial Societies and Social Simulation* 16 (4) 13.

[29] Watson, J. D., & Crick, F. H. C. (1953), A structure for deoxyribose nucleic acid. *Nature* 171, 737–738.