Choosing the Right Model — Robert Marks

Modelling
(from March & Lave (1975))

A. **What is a model?**

B. **Why model?**

C. **What is a good model?**
   (Any model, not just an ABM.)

A. **A model:**
   - **a simplified picture of a part of the real world.**
   - **has some of the real world's attributes, but not all.**
   - **a picture simpler than reality.**

We construct models in order to explain and understand.

Three Rules of Thumb for Model Building:

1. Think "process".

2. Develop interesting implications.

3. Look for generality/robustness.

Judge models using: truth, beauty, justice.

That is, an interplay between the real world (truth), world of æsthetics (beauty), world of ethics (justice), and the model world.

< >

Rule: "stop having kids when your sons outnumber your daughters"

"Question: (Are you ready?) What will be the ratio of boys (with) to girls (without) in such a society?"

A Surprise —

$\rightarrow$ **for *most couples*: more sons than daughters.**

  but —

  for *society*: more girls than boys, almost always.

Let's simulate this using NetLogo.

http://www.agsm.edu.au/bobm/teaching/SimSS/NetLogo41-models/boysngirls2.html

< >

All Models Require Assumptions

Closed-form models often require assumptions for tractability (so that the modeller can solve the mathematical problem).

Simulation models also require assumptions (to simplify reality).

But the simulation model can be made much closer to reality than many closed-form models.

The trade-off: the exact answer to the wrong question (closed-form), or an approximate answer to the right question (simulation models)
— John Tukey (1915−2000)

Need measures of closeness to reality, for sets of time series.

< >

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.

Perhaps code using a different platform, or dock.

< >

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? (i.e. description, positive)**
- **what could be? (i.e. existence, plausibility)**
- **what should be? (i.e. prescription, normative)**

< >

Validation

Ideally: compare the simulation output with the real world.

But:

1. *stochastic* $\therefore$ complete accord is unlikely, and the distribution of differences is usually unknown

2. *path-dependence*: output is sensitive to initial conditions/parameters

3. test for "retrodiction": reversing time in the simulation; or: test from a past date to the present: calibrate with history

4. what if the model is correct, but the input data are bad?

Use Sensitivity Analysis, to ask:

-   robustness of the model to assumptions made
-   which are the crucial initial conditions/parameters?

use: randomised Monte Carlo, with many runs.
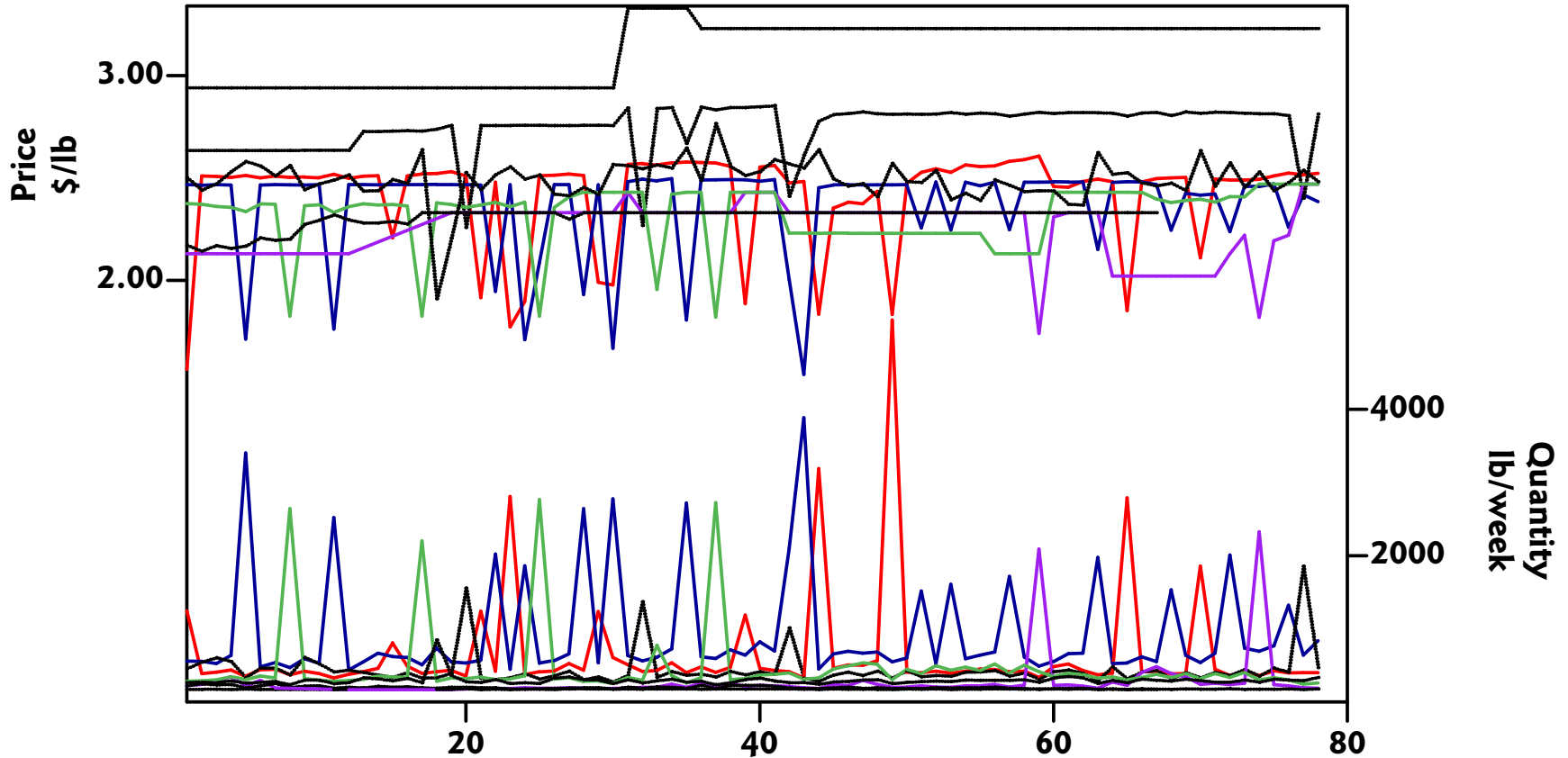
< >

Consider the following historical market data:

Figure 1: Weekly Prices and Sales (Source: Midgley et al. 1997)
(Coloured lines: Folgers, Maxwell House, Hills Bros, CFON)

Stylised Facts of the Market Behaviour

- **Much movement in prices and quantities of four brands — a rivalrous dance.**

- **Pattern: high price (and low quantity) punctuated by low price (and high quantity).**

- **Another four brands: stable prices and quantities**

Questions:
What is the cause of these patterns?

- **— shifts in brand demand?**

- **— reactions by brands?**

- **— actions by the supermarket chain?**

- **— unobserved marketing actions?**

< >

Explanations?

Interactions of profit-maximising agents, plus external or internal factors $\rightarrow$ via a model $\rightarrow$ behaviour

Similar (qualitatively or quantitatively) to the brands' behaviours of pricing and sales.

Note: assuming profit-maximising (or purposeful) agents means that we are not simply curve-fitting or description using D.E.s. Going beyond the rivalrous dance.

< >

Validation and Measurement

Q: how can we measure the degree of similarity of two sets of time-series?

***One*: the historical record of the rivalrous dance among the sellers in an oligopoly, while**

***The other*: the output from a (agent-based) simulation model of the market, where each seller agent prices this week as a function of the state of the market last week (or earlier).**

Q: how can we *output validate* our model against history?

Or: how can we derive a *degree of confidence* in the model output?

< >

The Issue: Heterogenous Agents and Time-series Prices

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 measure how closely the output reflects history, to *validate* the model.**

We are interested in the second, having used machine learning (the Genetic Algorithm) to derive 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).
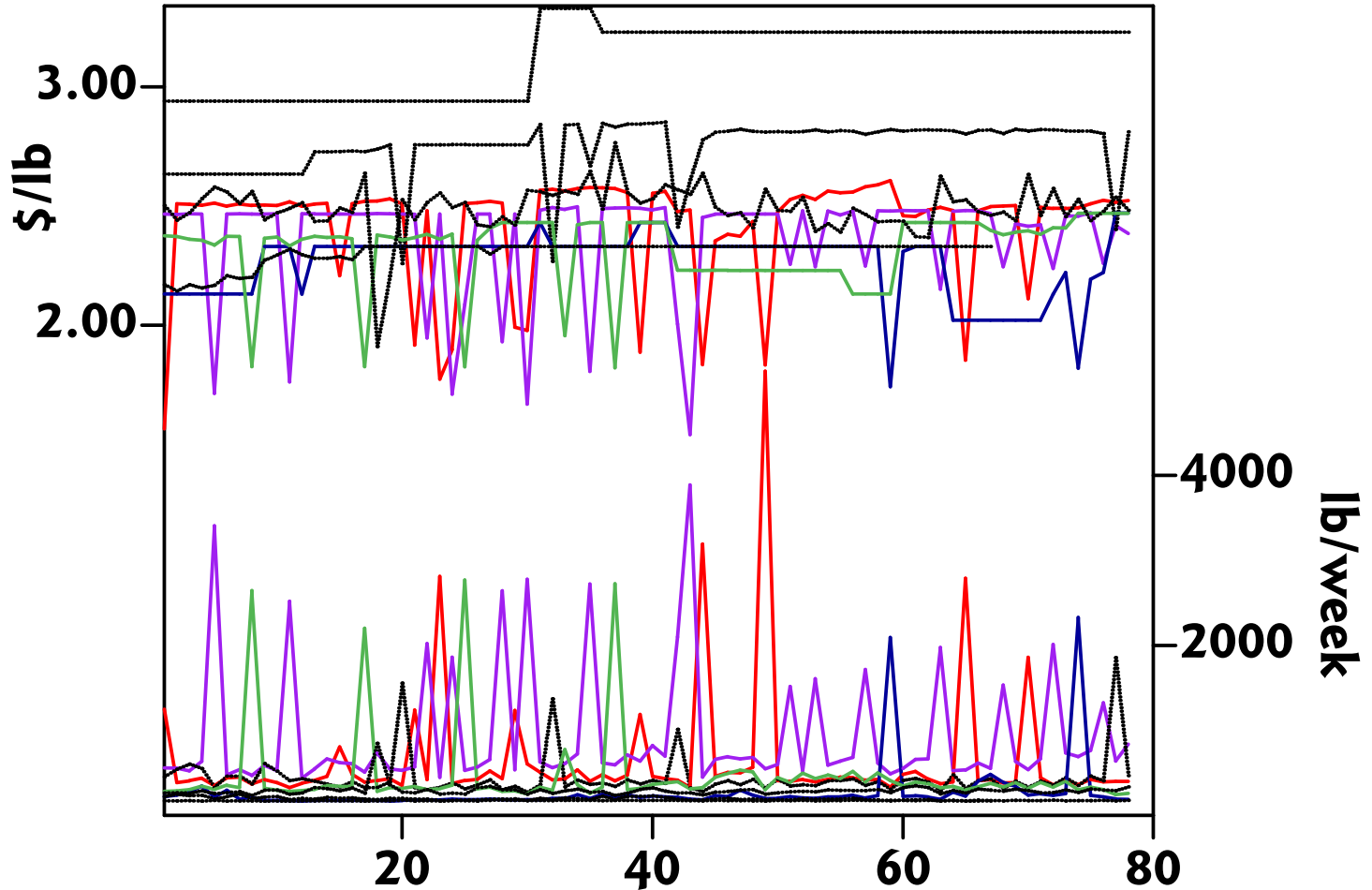
< >

Historical Data: Market Prices and Volumes

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

Dichotomous Price Partitioning of the Historical Data
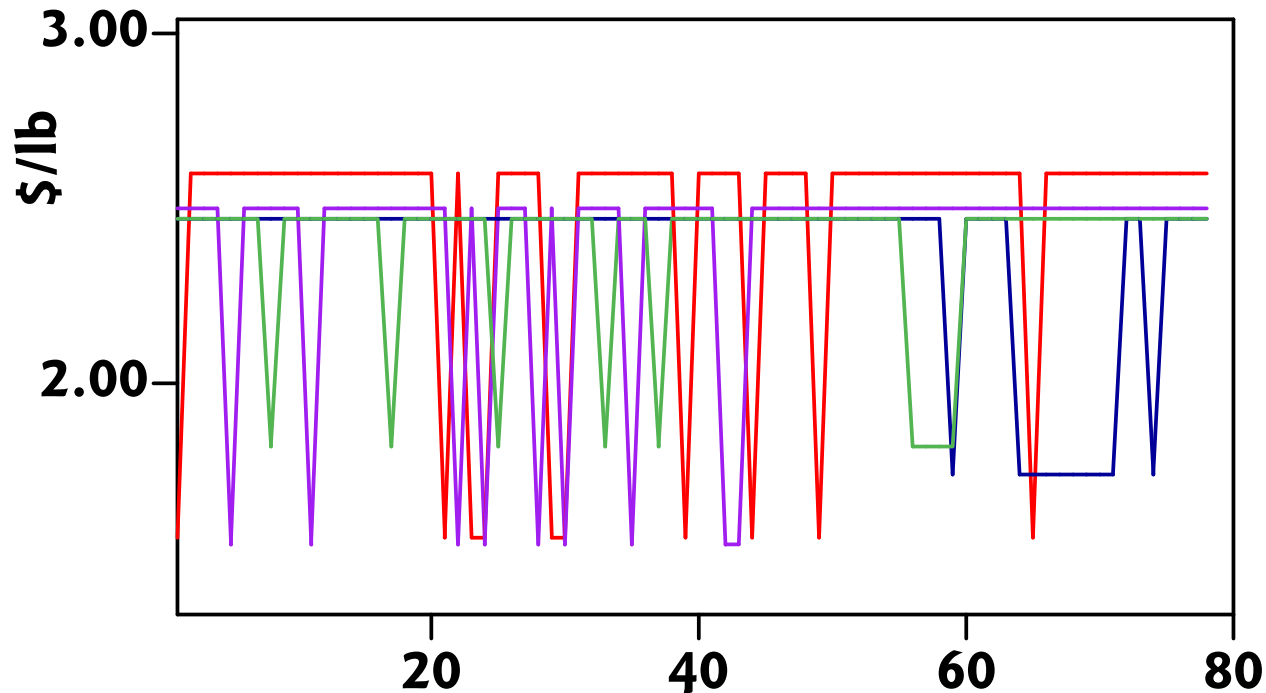
To handle the curse of dimensionality.

Figure 2: **Partitioned Historical Weekly Prices of the Four 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 at week w, where M_w in turn is the product of the weekly prices S_w of all brands over several weeks:

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

Earlier in the research program undertaken with David 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 model, and derived its simulated behaviour, as time-series patterns (below).

< >

The State Similarity Measure (SSM)

The SSM derives the distance between two sets of time-series, by calculating the sum of absolute differences in observed window states between the two set, 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 (below).

< >

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.
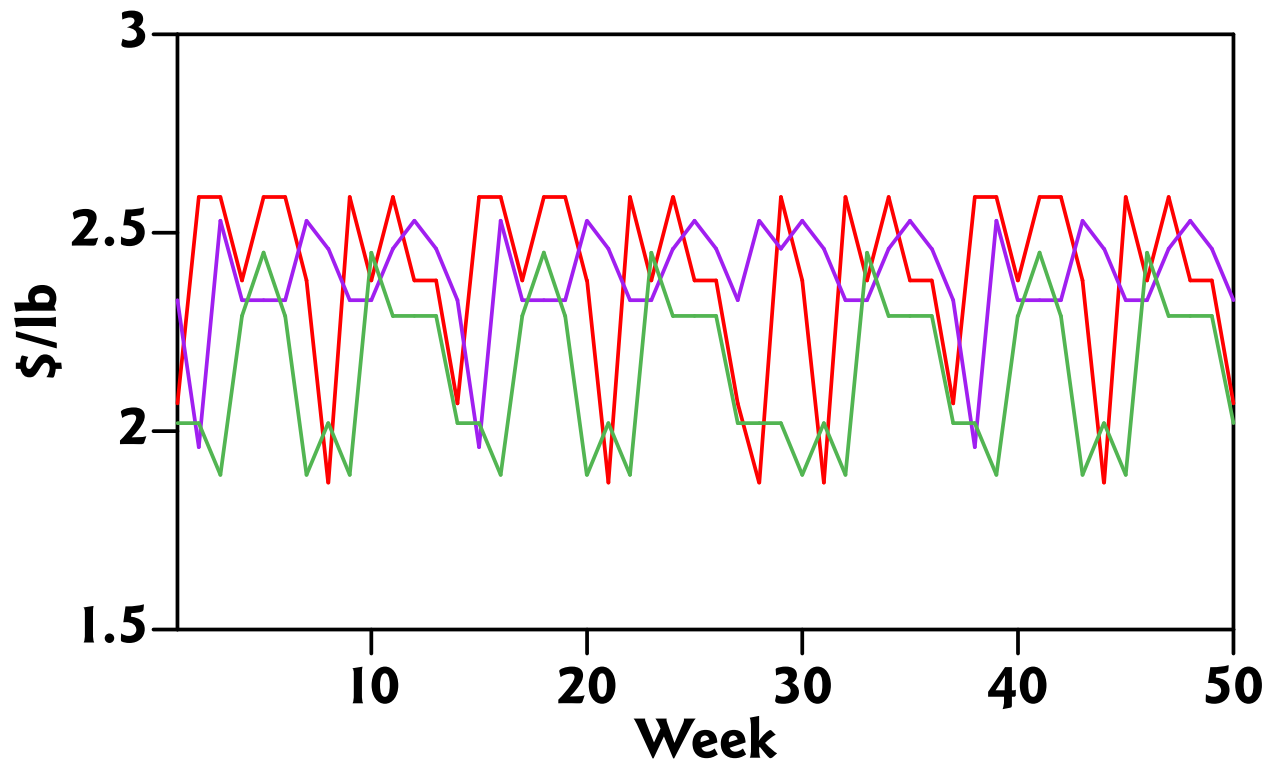

*Figure* 3: **Example of a Simulated Oligopoly (Marks et al. 1995)**

< >

Distances Between History and Three Runs (Brands 1, 2, 5)

	History	Run 11	Run 26a	Run 26b
History	0	82*	68	68
Run 11	82*	0	66	60
Run 26a	68	66	0	30
Run 26b	68	60	30	0

Table: **Distances Between History and Three Runs (Brands 1, 2, 5)
(* : cannot reject the null at the 5% level)**

Here, S, the maximum number of states = 48, so the maximum distance apart is 96. The three Runs are closer to each other than to the Historical Data; Runs 26a and 26b are very close, only 30/96 = 31.25% apart.

< >

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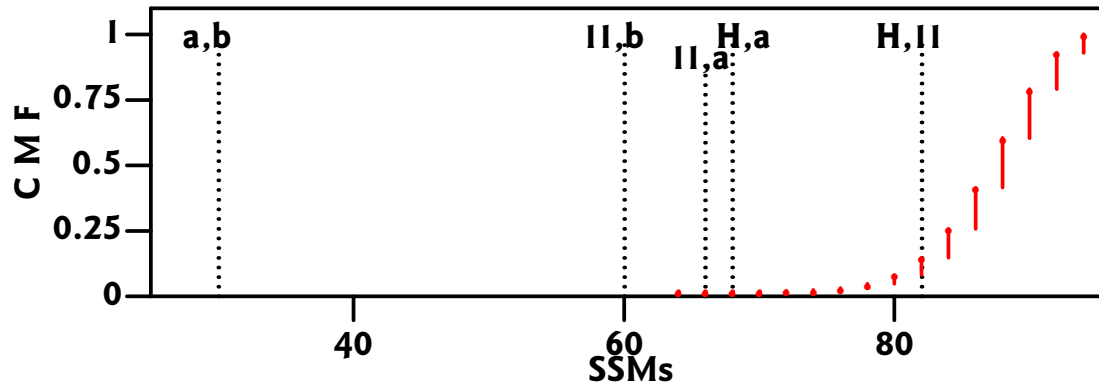
The red lines are the CMF of pairs of sets of random series (3 series, 48 observations) from 100,000 Monte Carlo parameter bootstraps.

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

Cannot reject the null hypothesis (random sets) for Historical data and Run 11; reject the null (random) hypothesis for all other pairs.

< >

Conclusions — the State Similarity Measure

This measure, *the State Similarity Measure* (SSM), is sufficient to allow us to put a number on the degree of similarity between two sets of time-series which embody dynamic responses.

Such a metric is necessary for scoring the distance between any two such sets, which previously was unavailable.

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

The SSM will also allow us to measure the distance between any two sets of time-series and so to estimate the parameters, or to help calibrate a model against history.

< >

References

[1]   March J. and Lave C. (1975), *Introduction to Models in the Social Sciences*, (New York: HarperCollins).

[2]   Marks R.E. (2010), "Comparing two sets of time-series: the State Similarity Measure," In *2010 Joint Statistical Meetings Proceedings — Statistics: A Key to Innovation in a Data-centric World*, Statistical Computing Section. (Alexandria, VA: American Statistical Association), pp. 539–551.
       http://www.agsm.edu.au/bobm/papers/asap.pdf

[3]   Marks R.E. (2012) "Monte Carlo," in *The Palgrave Encyclopædia of Strategic Management*, edited by David Teece and Mie Augier, (London: Palgrave), forthcoming.
       http://www.agsm.edu.au/bobm/papers/MonteCarloPESM-longer.pdf

[4]   Marks R.E., Midgley D.F., and Cooper L.G. (1995) "Adaptive behavior in an oligopoly," *Evolutionary Algorithms in Management Applications*, ed. by J. Biethahn and V. Nissen, (Berlin: Springer-Verlag), pp.225–239.

[5]   Midgley, D.F., Marks, R.E., and Cooper, L.G. (1997) "Breeding competitive strategies," *Management Science*, 43(3): 257–275.
       http://www.santafe.edu/media/workingpapers/95-06-052.pdf