

Detecting Key Variables in System Dynamics Modelling by Using Social Network Metrics

J. Barranquero, M. Chica, O. Cerdón, and S. Damas

1 Introduction

Strategic management requires special economic and marketing planning, offering the ability to comprehend and anticipate the effects of complex dynamic interactions between a firm and their business environments and stakeholders. This is therefore a complex social system that requires understanding emergent patterns and their systemic implications (Bonabeau 2002; Dickson et al. 2001). A concrete example of this strategic problem is the brand value management, where decision makers must consider the outcome of their investments to make a sustainable and differential advantage relative to their competitors (Aaker 1996).

Building a business dynamic model that lays out the critical resources of the scenario and the key relationships between them offers a competitive gain for decision makers. This kind of modelling also provides a way to carry out simulations and understand the effects of the different policies. Among other methodologies, system dynamics (SD) (Forrester 1961; Sterman 2000) presents a theoretical framework with a set of tools and techniques for developing mathematical models of complex systems for social and economic scenarios.

The SD methodology is particularly useful in systems with many interrelated variables, where relevant data to build the system is not always available. SD offers

J. Barranquero • M. Chica (✉) • S. Damas
European Centre for Soft Computing, Mieres (Asturias), Spain
e-mail: jose.barranquero@softcomputing.es; manuel.chica@softcomputing.es;
sergio.damas@softcomputing.es

O. Cerdón
European Centre for Soft Computing, Mieres (Asturias), Spain
DECSAI and CITIC-UGR, University of Granada, 18071 Granada, Spain
e-mail: oscar.cordon@softcomputing.es; ocordon@decsai.ugr.es

the opportunity to simulate a problem by investigating its results and behaviour, making the framework useful for policy testing, what-if scenarios, or policy optimization.

The set of applications of SD is enormous (O'Regan and Moles 2006; Winz et al. 2009). Besides, it has played an important role for a systemic view of management issues (Warren 2005) and marketing applications. The application of SD for brand management assists marketing experts in understanding how different factors affect the value of a particular brand, how costumers react to a brand in terms of loyalty and equity, the influence of email marketing campaigns, or the effects of implementing innovation policies in organizational policies (Mukherjee and Roy 2006; Richardson and Otto 2008).

However, it is sometimes difficult to identify key variables in dense or large problems modelled by SD. These key variables are those able to generate significant changes in the whole system. This descriptive information of the system is vital for modellers since they can apply strategic actions over those variables (in a direct or indirect way) and focus their what-if scenarios. The identification of these key variables is also useful for understanding the dynamics of the model and for validation purposes, given that key variables might constitute an additional boundary adequacy and structure verification test for the model (Oliva 2003; Quadrat-Ullah 2012).

The main focus of the current paper is how to detect which variables of a SD model constitute the set of key variables. Our proposal is to first compute a quality metric for every variable of the graph structure of the model. These values indicate the importance role of each variable with respect to the whole structure of the model. Then, we rank model's variables according to this metric, suggesting those that yield better values.

Hence, our proposed quality metric is founded on network-based properties of the model structure and is therefore applied on the whole SD graph. The computation of the metric is based in turn on the *scope* and *closeness* of an agent within a social network, which are well known metrics in social network analysis (SNA) (Carrington et al. 2010; de Nooy et al. 2005; Oliveira and Gama 2012).

We have modelled and simulated a TV show brand management problem to validate the application of our key variable detection algorithm. This systemic abstraction is based on an existing work that analyzes the Indian version of "Who wants to be a millionaire" (Mukherjee and Roy 2006). We have followed Vester's sensitivity model (Vester 1988, 2007) to shape the system dynamics and structure. This SD methodology is convenient for sustainable processes and enables analysts to simplify the real world complexity into a simulation and consensus system. After applying the algorithm and extracting the key variables of the model structure we run different simulations to compare the global impact of injecting strategic actions just over top-ranked key variables.

The rest of the paper is structured as follows. In Sect. 2 we study the background and describe the SD modelling and social networks metrics of our proposal. Then, Sect. 3 contains the analysis of results and simulation graphs of the key variables

detection for the TV show case. Finally, we present some concluding remarks in section Concluding Remarks.

2 Methods

2.1 System Dynamics for Modelling Complex Marketing Systems

There are different methodologies and tools for system dynamics modelling. For our work we follow the sensitivity model proposed by Vester (1988, 2007) which offers a semi-quantitative SD modelling tool based on systems thinking and fuzzy logic (Zadeh 1975). It has been applied to different fields of research, environmental and risk management, and tourism (Huang et al. 2009; Meyer-Cech and Berger 2009; Schianetz and Kavanagh 2008). The main advantages of this approach are the ease of use and the employ of feedback analysis as the core component of the modelling process.

There are nine steps in sensitivity modelling. These include system description, set of variables, criteria matrix, impact matrix, systemic role, effect system, partial scenarios, simulation and cybernetic evaluation. They can be categorized into three phases (see Fig. 1). The first phase begins with a general system description and the

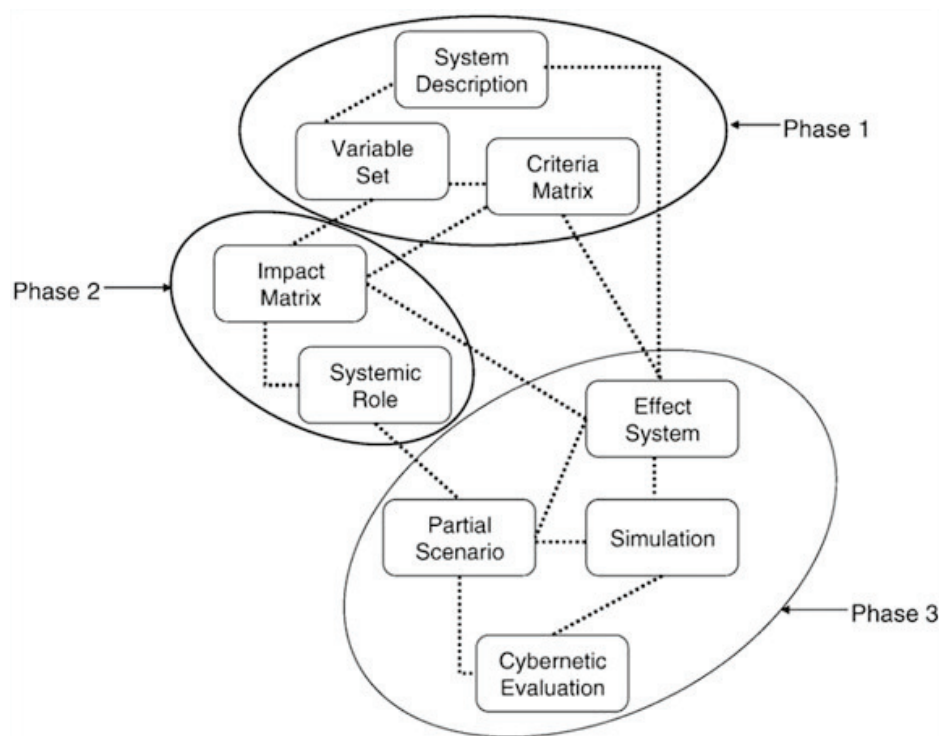


Fig. 1 The main steps of the sensitivity model methodology (Huang et al. 2009)

identification of influential factors for system development. In the second phase there is an analysis of the magnitude of the cause-effect relations between the variables and to identify their functional roles in the system (model structure). Simulation and cybernetic evaluation of phase 3 are based on the framework of the effect system and the positive and negative feedback relationships. The partial scenario of the focus issue can be simulated to observe the dynamics and inter-relationships among variables.

One of the major components of this SD methodology is the feedback loops that embody the information feedback structure of the system. Feedbacks are defined by the effects among the variables of the system. An effect between two variables can be direct or inverse. The effects, in conjunction with the variables, form the graph structure of the system. See Fig. 2 for a structure example of the case study of this work.

The simulation results arise from this interaction among feedback loops. Feedbacks are of two types: mitigating (an initial change in one variable of the loop will finally change the variable in the opposite direction, balancing the initial change) and reinforcing (where the initial change will be reinforced through the feedback process).



Fig. 2 System model structure for TV show case study

For this work, we extend the original SD methodology to tackle economical and marketing problems as ours. Therefore, the modeller can also define the details of effects (delay, intensity, and values of change) and variables (initial and optimum value, limits, blocked status, and randomness). In order to run simulations the model allows the definition of the temporal horizon and different strategies to be applied to some action variables to optimize the state variables of interest.

2.2 *Use of Social Network Analysis for Key Variable Detection*

We define the key variables of a system as those inherently relevant due to their interconnection with others variables of that system. We consider that a key variable is not required to be a good lever for defining specific actions, even when it plays an important role in the development of the system.

Therefore, instead of focusing on specific simulation conditions for key variable detection, we consider that it is more interesting to emphasize the intrinsic structure of the network, which in fact represents the system dynamics through its effects. This shall generate a more stable and general key variable set, suitable for a broader range of configurations.

2.2.1 Social Network Centrality Metrics

Centrality metrics are typically applied in SNA, making intensive use of statistical graph-based measures (Carrington et al. 2010; de Nooy et al. 2005; Oliveira and Gama 2012). SNA distinguish two levels of analysis: individual units (variables, actors, etc.) and whole network. In this paper, we focus on the former, given the definition of key variables that we want to address, clearly oriented to explore the role of each individual. Whole network indicators may be also helpful for obtaining more condensed knowledge, although we do not cover them in detail for this work due to length constraints.

Therefore individual SNA metrics can be adapted for our purposes, given that they share a common objective of identifying key actors in a network. Globally, these metrics are considered as a measure of *centrality* or *prestige*, in which the most common (Bonacich 1987; Freeman 1979) are:

- **Degree or valency:** analyzes the immediate neighbourhood of each node and is computed as the number of edges of the node. For directed graphs it is divided into in-degree (input prestige or support) and out-degree (output prestige or influence). Its biggest drawback is that it is a local measurement that does not reflect the global structure of the network.
- **Betweenness:** measures the relevance in terms of the number of shortest paths that go through the node, although it can be also computed for edges. Nodes (or edges) with high betweenness are supposed to interact heavily in information

flow and diffusion between communities (i.e., to play a strong brokerage in the information diffusion process).

- **Closeness**: evaluates how fast a given individual can reach the whole network, defined as the average length of all shortest paths with origin on the node. Its main problem is that it is not defined for those cases where there exist pairs of nodes that are not connected by any path.
- **Eigenvector centrality**: a re-elaboration of degree, which takes into account the quality of first order connections. It is computed after assigning a relative score to each node, measuring their connectivity with other well-connected nodes.

2.2.2 The Proposed Quality Metric

The reviewed metrics lack to take into account the number of reachable nodes, which can be defined as its *scope*. The scope measures how many variables can be reached directly or by transitivity. For instance, the standard closeness, defined in the previous section, could be highly skewed. This is because variables with low scope may show optimal values of closeness, while being poorly connected. In other words, it is considered a local measure.

Therefore, we have adapted the original definition of closeness in order to take into account the scope and the delay of each effect (weight of the edge in our model). This proposal allows measuring both concepts jointly, considering on the one hand the elasticity and penalizing on the other hand the absence of a path between nodes.

The new metric is termed *elastic distance* (ED) and is computed as the average of the shortest weighted distances from the source variable to all other variables. The distances to non-reachable ones are fixed with a sufficiently large value in order to embody information about the scope of the variable under study, defined as supreme-distance constant M :

$$ED(i) = \frac{1}{n} \sum_{j=1}^n d(i, j), \forall i \neq j, \quad (1)$$

where $d(i, j) = M$ when there is not any path between nodes i and j .

3 Experimentation Results for a TV Show Case Study

In this section we apply our proposed metric to a SD model adapted from the original model of the Indian “Who wants to be a millionaire” TV show (Mukherjee and Roy 2006). The graph of the model is presented in Fig. 2. Node colours depend on current value of corresponding variables, while node diameter is defined in terms of number of feedbacks in which the variable participates. The name, metric value

Table 1 Metric values obtained for variables defined in TV show case study

ID	Name	Metric	Description
12	Interest level	5.93	Perceived interest by show viewers. Equivalent to brand equity defined by Aaker (1996)
2	Innovation	6.20	Action variable that allows influencing the system by simulating novelties like special editions
7	Brand loyalty	6.33	State variable that measures actual viewers loyalty. One of the measures of brand value defined by Aaker (1996)
14	Actual viewers	6.47	Measures the success of the show in terms of total number of viewers. Critical state variable
13	Host popularity	6.80	Current popularity of the person driving the show. Reinforcing feedback cycle with interest level
1	Repetitiveness	6.87	Measures the degree of repetitiveness of the show. It is directly influenced by episodes rate and innovation
9	Probability of joining	7.27	Probability of viewers joining in. It depends on interest level and potential viewers
5	Episodes rate	7.40	Number of episodes per time unit. Too many episodes affect negatively to repetitiveness and brand loyalty
6	Potential viewers	8.07	Available viewers that do not follow the show currently. Inversely related with actual viewers
10	Minimal promotion level	8.73	Minimum spending for promoting the show. This variable tends to increase over time
11	Promotion effectiveness	8.87	Effectiveness of investment in promoting the show. It acts directly over interest level
15	Channel popularity	9.47	Similar to host popularity, though less influenced by show's interest level
16	Competition	9.53	Amount of competing shows. Influenced by actual viewers and influences brand loyalty
8	Promotion expenditure	10.73	Amount of money spent on promotional campaigns. Influenced by advertisement revenue
3	Advertisement rate	13.60	Advertisement benefit per time unit. It is heavily influenced by actual viewers
4	Advertisement revenue	13.67	Advertisement incomes per time unit. Closely related with advertisement rate

and description of each variable is detailed in Table 1 (variable list is sorted in terms of elastic distance).

3.1 Results of Key Variable Detection Algorithm

The two highest ranked variables are *interest level* (5.93) and *innovation* (6.20), while variables with the lowest rank are *advertisement rate* (13.60) and *advertisement revenue* (13.67). Given that our proposed metric is a weighted distance, this ranking reflects the speed and scope of propagation of changes for each variable.

Interest level is ranked first because it has many outgoing paths, spreading changes over the rest of system variables very quickly. The case of *innovation* is quite similar, because it allows a rapid access to core system variables, including interest level.

Advertisement rate and *advertisement revenue* obtain bad ranks because they are placed at the beginning of a long path, which implies a slower propagation over the whole network.

3.2 Simulation Results Using the Set of Key Variables

In order to validate our proposed algorithm for key variable detection, we present two simulation scenarios in Fig. 3. The idea is to test if there exists a significant difference between acting over top ranked versus bottom ranked variables.

The system is initially configured to be in a relatively steady state, avoiding strong trends that could clutter the interpretation of simulation results, with respect to a simulation baseline. The simulation engine evolves all variables of the system with a range that goes from 0 to 100, representing abstract values without specific representation units.

We simulate two alternative strategies with two actions each. The first strategy (Fig. 3, left graph) acts over the two lowest ranked variables, modifying their values to optimum. We refer to this scenario as *strategy L*. The second one (Fig. 3, right graph) applies the same change to optimum values over the two highest ranked variables, referred as *strategy H*. Finally we also simulate the system without any action, in order to measure the % of change of each variable with respect to *baseline simulation*. All simulations are performed over a 1 year period.

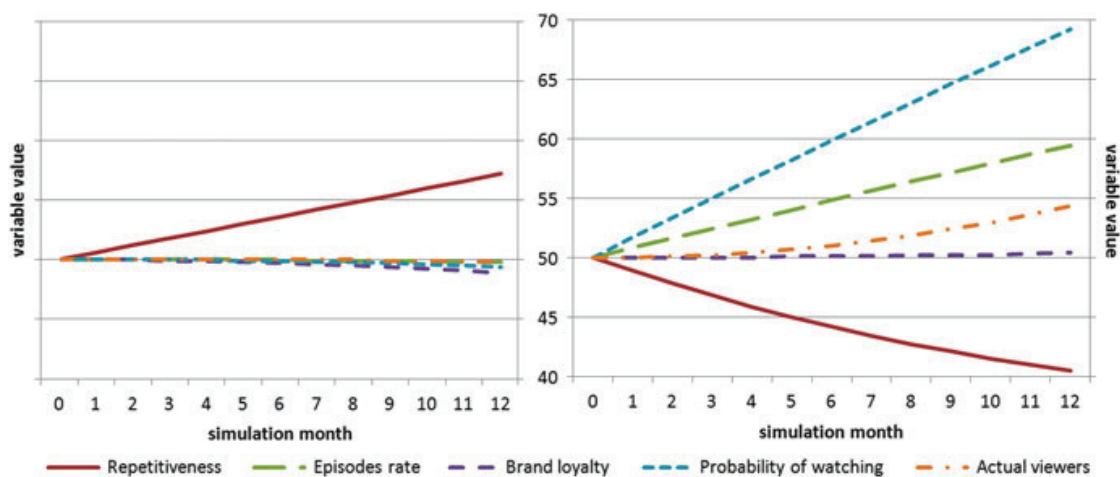


Fig. 3 Evolution of the system when directly acting over the lowest ranked variables (strategy L, left) and over the top key variables (strategy H, right)

We analyze the obtained simulation results excluding actioned variables from both strategies and measuring absolute changes with respect to baseline simulation. We also define two basic metrics, *total absolute distance* (TAD):

$$TAD = \sum_{i=1}^n |v_i - v'_i|, \quad (2)$$

and *mean absolute distance* (MAD):

$$MAD = \frac{1}{n} \sum_{i=1}^n |v_i - v'_i|; \quad (3)$$

where v_i and v'_i represent the final value of variable i for simulated baseline and strategy respectively.

Strategy L only affects 25 % of variables (3/12), obtaining a TAD of 15.6 and a MAD of 1.3. On the other hand, strategy H produces changes in 67 % of variables (8/12), with a TAD of 71.5 and a MAD of 6.

The highest change produced by strategy L is achieved over *promotion expenditure*, with a variation of 14.8 (out of 15.6 TAD). This is because it is the closest variable on the main path that starts on actioned variables. In this simple case it may seem obvious that the strategy is worthless, but at least the experiment succeeds in validating that the algorithm is effectively detecting the less critical variables.

Strategy H achieves an outstanding improvement with respect to both the baseline and strategy L. The strongest changes are produced over *probability of watching* (19.8), *host popularity* (18.9), *repetitiveness* (16.7), and *episodes rate* (9.6). All of these variables are in turn on the middle upper part of the ranking of key variables, producing a snow ball effect that shall be more significant for longer simulations.

Concluding Remarks

A key variable detection algorithm to be applied over the structure of a SD model was presented in this work. A quality metric is calculated for each variable of the model to quantify its importance for changing the evolution of the system. This metric is an extension of closeness, a widespread SNA measure, that we use for ranking the set of variables of the model structure.

The results of the proposed detection algorithm can be an effective validation test for the designed model. The set of key variables point out which variables are prevailing in terms of the model structure. Hence, if the set does not fit with the intended idea of the system, probably the design of the model may require a revision.

We tackled a brand management problem for a TV show, modelling it by SD and applying the key variable detection algorithm. Results showed how

(continued)

variables *interest level* and *innovation* were the most important of the model yielding a metric value of 5.93 and 6.20, respectively. Strategic actions were applied to these variables to present the impact in the simulation results with respect to a baseline simulation. Our experiments showed that acting over these two key variables have a remarkable effect over system variables, with an average improvement of 6 points (MAD metric) over system variables in 1 year. The same experiment but applied over the two lowest ranked variables only produced an average deviation of 1.3 points (MAD metric).

Some future works arise from this contribution: (1) propose and evaluate other quality metrics such as *Local Clustering Coefficient* (Watts and Strogatz 1998) or algorithms like *Pagerank* (Brin and Page 1998; Easley and Kleinberg 2010); (2) include additional measurements for providing complementary information about the model structure; and (3) develop an optimization engine for defining appropriate strategic actions in order to maximize profitability in what-if scenarios.

Acknowledgements This work has been supported by RØD Brand Consultants under ZIO project and Ministerio de Economía y Competitividad under SOCOVIFI2 (TIN2012-38525-C02-01 and TIN2012-38525-C02-02) projects.

References

- Aaker DA (1996) Measuring brand equity across products and markets. *Calif Manag Rev* 38(3):102–120
- Bonabeau E (2002) Predicting the unpredictable. *Harv Bus Rev* 80(3):109–116
- Bonacich P (1987) Power and centrality: a family of measures. *Am J Sociol* 92(5):1170–1182
- Brin S, Page L (1998) The anatomy of a large-scale hypertextual Web search engine. *Comput Netw ISDN Syst* 30(1):107–117
- Carrington PJ, Scott J, Wasserman S (2010) *Models and methods in social network analysis*. Cambridge University Press, Cambridge, UK
- de Nooy W, Mrvar A, Batagelj V (2005) *Exploratory social network analysis with Pajek*. Cambridge University Press, Cambridge, UK
- Dickson PR, Farris PW, Verbeke WJ (2001) Dynamic strategic thinking. *J Acad Mark Sci* 29(3):216–237
- Easley D, Kleinberg J (2010) *Networks, crowds, and markets: reasoning about a highly connected world*. Cambridge University Press, Cambridge, UK
- Forrester JW (1961) *Industrial dynamics*. Cambridge, US
- Freeman LC (1979) Centrality in social networks conceptual clarification. *Social Netw* 1(3):215–239
- Huang SL, Yeh CT, Budd WW, Chen LL (2009) A sensitivity model (SM) approach to analyze urban development in Taiwan based on sustainability indicators. *Environ Impact Assess Rev* 29(2):116–125
- Meyer-Cech K, Berger H (2009) Spatial impact of a factory outlet center in a small Austrian community - the case study of Leoville. *disP Plan Rev* 45(176):19–30

- Mukherjee A, Roy R (2006) A system dynamic model of management of a television game show. *J Model Manag* 1(2):95–115
- Oliva R (2003) Model calibration as a testing strategy for system dynamics models. *Eur J Oper Res* 151(3):552–568
- Oliveira M, Gama J (2012) An overview of social network analysis. *Wiley Interdiscip Rev Data Min Knowl Discov* 2(2):99–115
- O'Regan B, Moles R (2006) Using system dynamics to model the interaction between environmental and economic factors in the mining industry. *J Clean Prod* 14(8):689–707
- Qudrat-Ullah H (2012) On the validation of system dynamics type simulation models. *Telecommun Syst* 51(2–3):159–166
- Richardson GP, Otto P (2008) Applications of system dynamics in marketing: Editorial. *J Bus Res* 61(11):1099–1101
- Schianetz K, Kavanagh L (2008) Sustainability indicators for tourism destinations: a complex adaptive systems approach using systemic indicator systems. *J Sustain Tourism* 16(6):601–628
- Sterman J (2000) *Business dynamics: systems thinking and modeling for a complex world*. McGraw-Hill, New York
- Vester F (1988) The biocybernetic approach as a basis for planning our environment. *Syst Pract* 1(4):399–413
- Vester F (2007) *The art of interconnected thinking: tools and concepts for a new approach to tackling complexity*. MCB Publishing House, Munich, Germany
- Warren K (2005) Improving strategic management with the fundamental principles of system dynamics. *Syst Dyn Rev* 21(4):329–350
- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393(6684):440–442
- Winz I, Brierley G, Trowsdale S (2009) The use of system dynamics simulation in water resources management. *Water Resour Manag* 23(7):1301–1323
- Zadeh LA (1975) Fuzzy logic and approximate reasoning. *Synthese* 30(3–4):407–428

Frédéric Amblard
Francisco J. Miguel
Adrien Blanchet
Benoit Gaudou *Editors*

Advances in Artificial Economics

Lecture Notes in Economics and Mathematical Systems

676

Founding Editors:

M. Beckmann
H.P. Künzi

Managing Editors:

Prof. Dr. G. Fandel
Fachbereich Wirtschaftswissenschaften
Fernuniversität Hagen
Hagen, Germany

Prof. Dr. W. Trockel
Murat Sertel Institute for Advanced Economic Research
Istanbul Bilgi University
Istanbul, Turkey

and

Institut für Mathematische Wirtschaftsforschung (IMW)
Universität Bielefeld
Bielefeld, Germany

Editorial Board:

H. Dawid, D. Dimitrov, A. Gerber, C-J. Haake, C. Hofmann, T. Pfeiffer,
R. Slowiński, W.H.M. Zijm

An Agent Based Propagation Model of Bank Failures	119
André Dias, Pedro Campos, and Paulo Garrido	
Direct vs. Side Effects in Financial Contagion: What Weights More?	131
Stefano Zedda	
Saudis and Expats: An Agent-Based Model of the Saudi Arabian Labor Market	139
Davoud Taghawi-Nejad	
Forbidding Fixed Duration Contracts: Unfolding the Opposing Consequences with a Multi-Agent Model of the French Labor Market	151
Olivier Goudet, Jean-Daniel Kant, and Gérard Ballot	
Shadow Economy and Wealth Distribution	169
Nuno Trindade Magessi and Luis Antunes	
Distribution Effects of Extortion Racket Systems	181
Klaus G. Troitzsch	
Impacts on Stability of Interdependencies Between Markets in a Cobweb Model	195
Emma Jonson, Liv Lundberg, and Kristian Lindgren	
Detecting Key Variables in System Dynamics Modelling by Using Social Network Metrics	207
J. Barranquero, M. Chica, O. Cerdón, and S. Damas	
Trade-In Programs in the Context of Technological Innovation with Herding	219
Paolo Pellizzari, Elena Sartori, and Marco Tolotti	
Evaluating Scenarios for Upgrading Sustainability of the Meat Supply Chain	231
Eva van den Broek and Tim Verwaart	