Intern. J. of Research in Marketing xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Agent-based modeling in marketing: Guidelines for rigor

William Rand [a,1], Roland T. Rust [b,*]

[a] Assistant Professor of Marketing, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, United States
[b] Distinguished University Professor and David Bruce Smith Chair in Marketing, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

Article history:
First received in March 26, 2010 and was under review for 5 months
Available online xxxx

Area Editor: Russell S. Winer

ABSTRACT

Agent-based modeling can illuminate how complex marketing phenomena emerge from simple decision rules. Marketing phenomena that are too complex for conventional analytical or empirical approaches can often be modeled using this approach. Agent-based modeling investigates aggregate phenomena by simulating the behavior of individual “agents,” such as consumers or organizations. Some useful examples of agent-based modeling have been published in marketing journals, but widespread acceptance of the agent-based modeling method and publication of this method in the highest-level marketing journals have been slowed by the lack of widely accepted standards of how to do agent-based modeling rigorously. We address this need by proposing guidelines for rigorous agent-based modeling. We demonstrate these guidelines, and the value of agent-based modeling for marketing research, through the use of an example. We use an agent-based modeling approach to replicate the Bass model of the diffusion of innovations, illustrating the use of the proposed guidelines to ensure the rigor of the analysis. We also show how extensions of the Bass model that would be difficult to carry out using traditional marketing research techniques are possible to implement using a rigorous agent-based approach.

1. Introduction

Marketing phenomena are often complex because they are the emergent result of many individual agents (e.g., consumers, sellers, distributors) whose motivations and actions combine so that even simple behavioral rules can result in surprising patterns. Moreover, these aggregate patterns feed back to affect individual choices. For example, consumers often make buying decisions based on their friends' advice or their social network, which affects product diffusion and influences the dominance of a brand in a market. However, the current dominant brand also affects an individual's decision as to which product to purchase. The diffusion pattern that results from the interaction of many consumers may in fact be much more complex than the adoption rules of the individuals. Thus, simple rules of behavior can give rise to complex, emergent patterns.

Agent-based modeling (ABM) is a tool that can help researchers understand and analyze these complex patterns (Epstein & Axtell, 1996; Gilbert & Troitzsch, 2005; Holland, 1995; LeBaron, 2000; Miller & Page, 2007). The basic concept of ABM is that by describing simple rules of behavior for individual *agents* and then aggregating these rules, researchers can model complex systems, such as the procurement of services in a marketplace, the purchase of tickets for events, or the adoption of innovations. An *agent* in an agent-based model is any autonomous entity with its own properties and behaviors; to develop an agent-based model, a researcher writes a description for each type of agent that details the agent's behaviors, properties, and the way the agent interacts with other agents and the environment. The power of ABM is that none of these descriptions requires knowledge of macro-dynamics; instead, the researcher encodes micro-rules of behavior and then measures the emergent macro-level results.

Despite the power of ABM, widespread acceptance and publication of this method in the highest-level journals has been slow. This is due in large part to the lack of commonly accepted standards of how to use ABM rigorously. Guidelines are needed for the proper use of ABM so that researchers, reviewers and editors who are unfamiliar with the methodology can still ascertain whether the approach was rigorously undertaken. In this paper, we address this need by proposing a set of guidelines for the rigorous development and analysis of agent-based models. It is important to establish these guidelines now because computational methods will become increasingly powerful and easier to implement as time goes on.

Thus, this paper is written in the spirit of the other methodological reviews often cited within the marketing literature that attempt to establish rigorous guidelines for a particular method. For instance, Anderson and Gerbing (1988) proposed guidelines for developing structural equation models along with a testing method to ensure that a model is as generalizable as possible. Similarly, Churchill (1979) and Gerbing and Anderson (1988) developed guidelines for the proper use

* Corresponding author. Tel.: +1 301 405 4300; fax: +1 301 405 0146.
E-mail addresses: wrand@rhsmith.umd.edu (W. Rand), rrust@rhsmith.umd.edu (R.T. Rust).
[1] Tel.: +1 301 405 7229; fax: +1 301 405 0146.

0167-8116/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2011.04.002

of scales. Similar to the approaches used in these papers, we not only discuss how ABM can be applied to marketing, but we also provide guidelines for how to rigorously apply ABM.

2. Previous applications of agent-based modeling in marketing

ABM has been used in a wide-range of fields of business-related research, from organizational science (Cohen, March, & Olsen, 1972) to supply-chain management (Walsh & Wellman, 1999). Several large firms, such as Procter & Gamble, have successfully applied ABM to improve revenues (North et al., 2010; Siebel & Kellam, 2003). Recently, there have also been special issues of the *Journal of Business Research* (Gilbert, Jager, Geffuant, & Adjali, 2007) and the *Journal of Product Innovation Management* (Garcia & Jager, 2011) focused on ABM.

With respect to marketing, ABM has often been applied to model the diffusion of innovations. For instance, several researchers (Goldenberg, Han, Lehmann, & Hong, 2009; Rahmandad & Sterman, 2008; Stephen, Dover, & Goldenberg, 2010; Watts, 2002; Watts & Dodds, 2007) have used an agent-based approach integrated with network science to model the role of influentials in diffusion. Goldenberg, Libai, Moldovan, and Muller (2007) and Goldenberg, Libai, and Muller (2010) have used ABM to explore network effects in product adoption. Garcia (2005) provides a good review of how ABM had been used in innovation research through 2005. Other researchers have used ABM to explore how both interaction networks and competitive forces affect firm positioning (Lusch & Tay, 2004; Tay & Lusch, 2002; 2005; Wilkinson & Young, 2002). Marks, Midgley and Cooper (1997) investigate adaptive firms using evolutionary algorithms in combination with an agent-based model (Marks, Midgley, & Cooper, 2006; Marks, Midgley, Cooper, & Shiraz, 1999), and Hill and Watkins (2007, 2009) have built upon established models (Axelrod, 1984) to examine moral behavior in marketing exchange relationships (Watkins & Hill, 2009).

A restricted form of ABM, known as cellular automata (CA), has also proven useful. For instance, Garber, Goldenberg, Libai, and Muller (2004); Goldenberg, Libai, Solomon, Jan, and Stauffer (2000); Goldenberg, Libai, and Muller (2001a,b); Libai, Muller, and Peres (2005, 2009); and Moldovan and Goldenberg (2003) use CA to model the diffusion of innovations. Goldenberg, Libai, and Muller (2002) use CA to examine cross-market communications, and Frels, Heisler, Reggia, and Schuetze (2006) use CA to examine network effects.

Toubia, Goldenberg, and Garcia (2008) use ABM to explore network effects on diffusion, as did Shaikh, Ragaswamy, and Balakrishnan (2005) and Delre, Jager, Bijmolt, and Janssen (2010). In this paper, we will also explore networks and agent-based models, but our work differs in that our goal is to establish general guidelines for the rigorous development and analysis of agent-based models. It should be noted that much of the previous work that has made its way into top tier journals has done so because it does use ABM in a rigorous manner; therefore, the goal of this current study is not to point to mistakes that have been made, but rather to establish guidelines that will help promote the expanded use of ABM in the future among researchers who otherwise might not know the most rigorous way to use ABM.

3. Why use agent-based modeling?

Before we describe the guidelines for implementing a rigorous agent-based model, we must take a step back and examine the reasons to use agent-based modeling. Because ABM models the individual, it can incorporate characteristics that are difficult to include in traditional models. For instance, consumers modeled with ABM can be boundedly rational (Arthur, 1994), heterogeneous in their properties and actions, adaptive and sensitive to history in their decisions (Rand, 2006), and located within social networks (Watts, 1999) or geographical locations (Brown, Riolo, Robinson, North, & Rand, 2005). ABM is most useful when the rules of behavior are easily written at the individual level; then the behavior of the system *emerges* (often referred to as an *emergent property* of the system).

Critiques of agent-based modeling often come from two points of view: one viewpoint is that ABM does not deal with real data and is therefore only for "toy problems", while another viewpoint is that most agent-based models have so many parameters that they can fit any data and are thus nothing more than "computer games." With regard to the first claim, it is definitely possible to create agent-based models that do not correspond to real-world phenomena, but as we will illustrate in this paper, ABM also provides a natural way to integrate real-world data and complexities into a model. Moreover, this critique is true of many modeling techniques. With any model, it is the researcher's burden of proof to show how the model corresponds to reality. We will examine how to ensure that an agent-based model is empirically valid in Section 7. As for the critique that ABM can fit any data set, this is not true if the model process, inputs and outputs are shown to be valid (i.e., they correspond to the real world). It is also possible to fit almost any data using regression if you are free to choose the functional form, but once the functional form has been selected and the dependent and independent variables have been chosen, there are rigorous methods by which to determine if the model fits the data. The same is true of ABM, and the goal of this paper is to explain these methods.

Before deciding to employ any modeling technique, it is important for the researcher to examine the strengths and weaknesses of different methods. Analytical modeling, for instance, gives researchers the ability to investigate the impact of various theories of firm and consumer behavior but can often be difficult to compare to real-world data and sometimes require assumptions that are overly simplistic. Empirical and statistical modeling, on the other hand, is a useful approach for describing extant data sets and making predictions about future statistical relationships, but such models rarely contain a theory of consumer behavior. Most consumer behavior experiments are useful for understanding the individual decisions that people make when confronted with marketing actions, but in most cases it is cost prohibitive or not possible to examine these effects on a large scale. System dynamics modeling (Sterman, 2000) requires the rules of behavior to be written at a higher level, such as how the whole population of consumers will respond to a marketing activity rather than how a particular individual will respond.

The strongest benefit of using an ABM approach within marketing is that the actions of firms and consumers within the model can be constructed based upon strong theories of behavior, but at the same time, the results can be validated against empirical data and the model can then be used to make predictions. The advantages and disadvantages of these various techniques, as well as a description of how ABM complements them, are summarized in Table 1.

As Table 1 shows, ABM is a natural complement to many other approaches. For instance, the trade-off between analytical modeling and ABM is explored in Fig. 1. The results of analytical modeling, unlike the results of ABM, are generalizable (i.e., the results are true for all parameters within the constraints of the assumptions). The results of ABM are only generalizable to the extent to which the model parameter space has been fully explored. However, analytical modeling often builds in simplifying assumptions about agent behavior (e.g., risk neutral, perfect information), while ABM allows researchers to relax many of these restrictive assumptions and build a model with more realism and complexity.

ABM is not just a solution to intractable analytical models. When agent-based models are properly validated, they add a layer of realism that is not captured by many analytical models. Moreover, the ABM model is built from the ground up to generate results, so instead of modeling aggregate patterns, those patterns are observed from a large number of individual decisions. Even when a closed form solution exists, an agent-based representation provides a

Please cite this article as: Rand, W., & Rust, R.T., Agent-based modeling in marketing: Guidelines for rigor, *Intern. J. of Research in Marketing* (2011), doi:10.1016/j.ijresmar.2011.04.002

Table 1
Comparison of marketing research methods.

Name	Advantage	Disadvantage	Complementary role of ABM
Analytical modeling	Generalizable, creates actionable insights into firm level strategic decisions	Difficult to compare to real-world data, sometimes requires overly simplistic assumptions	Agent-based models can be built from analytical models that include more realistic assumptions and can be compared to real-world data.
Empirical modeling and statistical modeling	Useful for finding patterns of behavior in extant data sets, and for making predictions about future behavior	Rarely linked to a behavioral theory at the level of the individual consumer or firm. Requires the right kind of data to exist showing relationships	If a theory of individual-level behavior can be generated, then agent-based models can be created that can be compared to empirical and statistical models.
Consumer behavior experiments	Provide theoretical insight into consumer decisions and reactions to marketing actions	Rarely scale up to large groups or examine complex consumer–consumer interactions	Agent-based models can be built upon consumer behavior theories and then scaled up to larger populations.
System dynamics modeling and other computational modeling forms	Allow a systematic examination of an entire complex system of interactions	Rules of behavior must be written at the system level and examination of individual-level heterogeneity can be difficult.	Agent-based models can complement larger scale models with a fine-grained resolution when necessary.
ABM	Allows the exploration of individual-level theories of behavior, but the results can be used to examine larger scale phenomenon.	Computationally intensive, not generalizable beyond the instances examined	

comparison to real-world data or an instantiation of consumer-level behavior theories.

When confronted with this trade-off, researchers can use both ABM and analytical modeling as complements and thus generate results that are both generalizable (if the restrictive assumptions hold) and built on a rich set of assumptions (if the restrictive assumptions do not hold). ABM works well within a multi-method context because the assumptions and parameters of ABM can be tied to the findings of other methods. This is important because it gives the researcher the ability to triangulate the models' findings. As Table 1 illustrates, ABM is useful in a context where behavioral rules, processes, and parameters are drawn from other methods. In these cases, ABM not only provides solutions to models that might have an intractable closed-form solution, but it also provides an additional level of realism.

More traditional analytical and regression models have often proven useful because they are computationally efficient.[2] However, as computational power becomes increasingly inexpensive, it becomes more efficient to employ computational modeling to understand complex systems. One way to think about this is as a usage problem with two substitutable goods (in this case, computational methods, such as ABM, vs. more traditional methods, such as analytical modeling) (Fig. 2). The cost of computation is steadily decreasing, while the costs of less computational methods (e.g., analytical modeling) are remaining constant.

Referring to Fig. 2, the line segment AB represents an isocost curve at time 1. In other words, each point on the line represents a combination of computational methods and analytical methods that costs the same as any other point on the curve. The researcher will seek to maximize utility, given a cost. The dotted lines show isoutility curves. Utility is maximized where AB is a tangent to the furthest possible isoutility curve. Thus, at time 1, the user chooses to use X_1 units of computational methods and Y_1 units of traditional methods.

Now consider time 2. The cost of computational methods has decreased, resulting in a new isocost curve AC. As a result, the user is now able to utilize more computation for the same cost. The user now chooses X_2 units of computational methods and Y_2 units of traditional methods. The relative mix has thus shifted in favor of computational methods ($X_2 > X_1$ and $Y_2 < Y_1$), which suggests that as computational costs decline over time, ABM and other computational methods will inevitably attract more use.

[2] A solution is computationally efficient if the time it takes to solve the problem scales as a polynomial or less with the size of the problem. By using assumptions to fix the input length of the problem, most analytical models are usually computationally efficient.

The benefit of this decreased computational cost has paid off in a number of research areas. As discussed in Section 2, ABM has been used substantially in the past to understand the diffusion of new products, for example, to explore the role of "hubs" (Goldenberg et al., 2009), but there are many other areas where ABM would be useful.

Table 2 illustrates a selected number of marketing application research areas that ABM is useful in exploring, along with example references of ABM research from marketing and non-marketing contexts. For instance, Heppenstall, Evans, and Birkin (2006) have constructed a model to explore retail location decisions using ABM. By combining ABM and geographic information systems (GIS), this model was uniquely able to show that small historical changes in local interactions can give rise to different global patterns; a phenomenon often referred to as path dependence. ABM has also been used in the past to examine inter-firm relationships, strategy and competition (Hill & Watkins, 2009; Lusch & Tay, 2004; Marks et al., 2006; Wilkinson & Young, 2002). ABM is useful in this application domain because as many firms as necessary can be modeled simultaneously and because the firms can incorporate complex learning models like genetic algorithms and q-learning, which facilitates the exploration of a richer space of firm strategies than is normally possible. This has allowed researchers to highlight the effects of those strategies on long-term competition and trust between firms.

Recently, there has been practitioner interest in using agent-based modeling to explore marketing mix models (North et al., 2010; ThinkVine, 2010). Though there has been some work in this area (Delre, Jager, Bijmolt, & Janssen, 2007), it is generally an area that could be more thoroughly explored using ABM. ABM is a powerful tool in this area because it has allowed researchers to include consumer-level

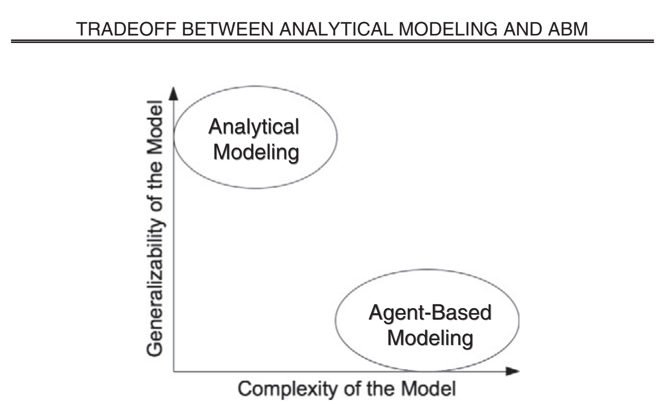

Fig. 1. Tradeoff between analytical modeling and ABM.

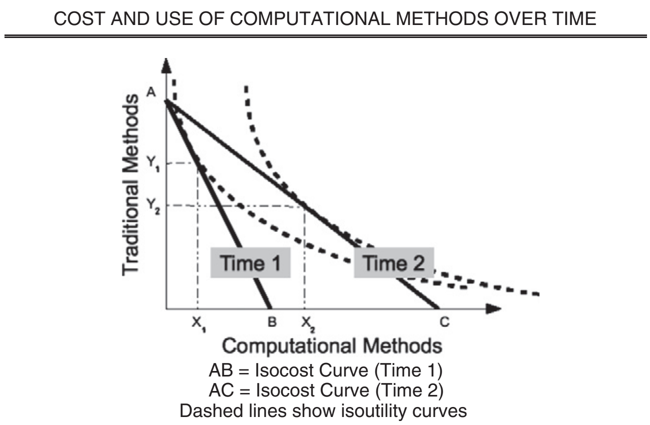

Fig. 2. Cost and use of computational methods over time.

behavior models of richer fidelity that have been used to examine the robustness of various marketing strategies. This research has produced substantial cost savings for at least one company (Procter & Gamble, in North et al., 2010). Finally, ABM could also be applied to retail and servicescape design. Though there is not yet any published research using ABM in this area, ABM would be useful because the movement patterns and behavioral decisions of each and every consumer can be modeled separately. Potential insights that could be gained in this area could relate to the examination of the interconnected role of service points and queuing, the ability to examine hundreds of thousands of retail location layouts in a small amount of time, and the discovery of unknown interactions between purchase decisions and consumer movement. The application of ABM to pedestrian modeling has proven useful in a number of other domain areas (Batty, 2005; Schelhorn, O'Sullivan, Haklay, & Thurstain-Goodwin, 1999; Torrens, in press), and this work could be applied within the context of marketing. These are just a few of the areas besides diffusion modeling where ABM could be applied within marketing.

4. Introduction to the illustrative example

Throughout this paper, we will illustrate the guidelines for rigorous use of ABM with one example. Our discussion of this example will be less detailed than usual due to space constraints and because it is only meant to be illustrative. Moreover, it is written from the perspective of the researcher who is developing a new agent-based model and not from the perspective of a final communication of the results of the model.[3]

The model that we will be developing is a model of consumer adoption—a version of the Bass (1969) model. The original Bass model of innovation diffusion was an aggregate model of diffusion; however, even in this aggregate model, two (p and q) of the three parameters (p, q, and m) are related to individual-level characteristics. The rates of adoption are based on mass media (p) and word-of-mouth (q). In the original Bass model, the decision to adopt at the population level is modeled as a hazard rate, but it would be interesting to examine how local social networks affect an individual's decision to adopt (Valente, 1995). To examine this idea, we first build an agent-based model that produces similar results to the original Bass model. We then investigate what happens in a network version of the model.

5. When is agent-based modeling appropriate?

Before we get to the model development itself, we should discuss when ABM is appropriate because this is really the first step in creating an agent-based model. The decision to use ABM should be based primarily on the question under investigation. If the question emphasizes groups of autonomous, heterogeneous entities that operate in a dynamic environment and if the measure of interest is an emergent result of these entities' interactions, then ABM is usually one of the tools that should be considered. Below, we present guidelines for when to apply ABM. However, these are not rules; instead they describe characteristics of problems that are amenable to using an ABM approach. These guidelines are specific to ABM and do not necessarily apply to other forms of computational modeling because they emphasize the individual-level modeling approach of ABM.[4] As we list these guidelines, we will specify whether they are indicative (the benefit of using ABM is increased if the problem exhibits this property), necessary (ABM is not appropriate if the problem does not exhibit this property) or sufficient (ABM is one of very few approaches that will work if the problem exhibits this property) for an ABM approach to be useful. The key indicators to consider in applying an ABM approach are the following:

1. Medium numbers (indicative) — ABM is not the appropriate tool to use when a system is composed of only one or two agents because, in that case, game theory often provides a better modeling tool. If the number of agents is very large and if the agents themselves can be modeled using a representative agent, then ABM becomes inefficient compared to statistical regression. The exact number of agents is relatively unimportant; medium numbers is a shorthand way of saying that though the system has a population of agents, this population can be affected by a few important individual interactions (Casti, 1995). Consumer adoption modeling, as in the Bass model, does exhibit the medium numbers property because most markets feature a group of consumers (i.e., influentials) that substantially affects the market's purchasing decisions.
2. Local and potentially complex interactions (indicative) — ABM becomes more useful as the interactions between individuals become more complex and local (Casti, 1995; Holland, 1995). Local information and complex interactions can be modeled using game theory, but often these models break down when the number of agents reaches above a small set. At this point, ABM becomes an appropriate framework to consider. Within consumer adoption, most consumers make local decisions based on their immediate social network, and the number of agents is large enough to warrant the use of ABM.
3. Heterogeneity (indicative) — Because the focus of ABM is on the individual, each individual can be modeled as differently from other individuals as necessary. For instance, agents can have different levels of willingness-to-buy, budgets, and demographic properties. Beyond different values, the agents can be of different types, such as consumers, companies, and the media. Types can even be divided into different types of organizations and consumers. For instance, risk-averse and risk-seeking organizations may appear identical according to their financial properties, but they will behave in different ways. Alternatively, if a system contains many homogenous agents, system dynamics modeling may be more useful because it efficiently tracks populations of identical agents and examines how they change over time. Within consumer adoption, most consumers have different local networks and different thresholds to innovation, which almost requires the use of an ABM approach.

[3] A standard protocol for publication of ABM results (Grimm et al., 2006; Parker et al., 2008) enhances communication and reproducibility (Polhill, Parker, Brown, & Grimm, 2008) but has been examined elsewhere.

[4] Computational models that also employ an individual-level modeling approach, such as some dynamic game simulations, can in fact be thought of as examples of ABM.

Table 2
Some marketing research application areas suitable for ABM.

Name	Advantage of ABM	Selected ABM marketing references	Relevant ABM references from other fields	Insights to gain/gained from ABM
Diffusion of information and innovations	Allows individual-level heterogeneity within both adoption decisions and social networks	Watts and Dodds (2007), Garcia (2005), Goldenberg et al. (2010), Stephen, Dover and Goldenberg (2010) and others in Section 2	Berger (2001) and Deffuant, Huet and Amblard (2005)	The structure of social networks and the role of hubs substantially affect diffusion.
Retail location decisions	Enables integration of individual-level behavior patterns with geographic information systems	Heppenstall et al. (2006)	Brown et al. (2005), Irwin and Geoghegan (2001) and Parker, Brown, Polhill, Deadman, and Manson (2008)	Local interactions can dramatically affect emergent outcomes (path dependence).
Inter-firm relationships, strategy and competition	Facilitates as many firms as necessary, each with firm-level characteristics, and the ability for firms to adapt their strategies over time	Hill and Watkins (2009), Wilkinson and Young (2002), Lusch and Tay (2004), Marks et al. (2006) and others in Section 2	Axelrod (1997), Barr and Saraceno (2002) and Prietula, Carley, and Gasser (1998)	The role of heterogeneous strategies in complex networks of firm relationships
Marketing mix models	Allows the examination of individual-level behavior patterns and reactions to the various elements of the marketing mix	Delre et al. (2007), ThinkVine (2010) and North and Macal (2010)	Specific to marketing but relevant simulations have been explored in supply networks: Parunak, Savit, and Riolo (1998)	Richer fidelity of models, incorporation of consumer-level behavior models, ability to test robustness of strategies
Retail and servicescape design	Can be used to model individuals moving about and making decisions in a complex retail environment	None as of yet	Batty (2005), Torrens (in press) and Schelhorn et al. (1999)	The interrelation of service points and queuing, examination of a large number of shop layouts, relationships between consumer movement and purchase behavior

4. Rich environments (indicative) — ABM facilitates the representation of rich and even dynamic environments (Gilbert, 2008). These environments can be as simple as two-dimensional abstract spaces or as realistic as a space derived from data contained in a geographic information system (GIS; Brown et al., 2005a) or a networked-based space derived with data from social network analysis (SNA; Carley, 2002). This allows ABM to capture the complexity of consumer adoption, from the influence of a convenience store's location on a consumer's decision to the effect of a consumer's peer network.
5. Temporal aspects (necessary) — ABM is a technique for modeling processes and is well suited for examining how complex systems change over time. Therefore, temporal aspects are almost a necessary condition for the ABM approach. Many modeling approaches allow you to examine the equilibrium states of dynamic games, but ABM is one of the few that allows you to examine the dynamics that give rise to those equilibria. To allocate resources within consumer adoption, it is often necessary to consider when people are likely to buy a product; therefore, the temporal nature of the process is central to the research question.
6. Adaptive agents (sufficient) — One of the promises of ABM is its ability to include adaptive agents within simulations (Holland, 1995). If an agent takes an action that produces a negative result, then that agent may try other actions in the future. An agent that changes its strategy (i.e., which actions to take in a given environment as a result of past information) is an adaptive agent. Because ABM is a computational method, it is possible to embed a machine learning approach within each agent that allows that agent to dynamically adopt the rules under which it operates (Rand, 2006). For instance, a retailer may alter its price setting and quantity-ordering actions, even when the state of the market is unchanged, due to previous experiences with different strategies (Marks, Midgley, & Cooper, 1997). There are few modeling techniques besides ABM that are able to robustly represent adaptation. Within consumer adoption, agents often make decisions to purchase a product based on previous recommendations from their social network. They dynamically adjust their confidence in their friends' opinions based on their own experiences with these products, and make future decisions based on this new trust network (Sharara, Rand, & Getoor, 2011).

6. Guidelines for model development

In general, building an agent-based model can be broken down into four large-scale steps: (1) decide if ABM is appropriate; (2) design the model; (3) construct the model; and (4) analyze the model.

Step 1. Decide if ABM is appropriate
Before building any agent-based model, it is important to first decide if ABM is in fact an appropriate approach. This is the premise of Section 5 and Table 1, where we explored the specific questions that should be analyzed to determine if ABM is appropriate and where we also determined that for our Bass model, ABM would be appropriate.

Step 2. Design the model
There are several decisions to be made when designing a model using the ABM approach (see Fig. 3):

(1) Scope of the model — What part of the complex system is the focus of the modeling effort and what aspects can be ignored? Within our current model, the scope will initially be defined based on the original Bass (1969) model. The model should be able to approximately reproduce the original results, and then it should be expandable to enable agents to possess a social network.

(2) Agents — The goal here is not to list every individual agent to be modeled, but rather to identify the general classes of agents and the approximate quantity of each agent type. The only agents in our current model are consumers.

(3) Properties — Each agent will need a list of properties that describe that agent. The properties of the agents in our model are initially based on the original macro-level model (Bass, 1969). Each agent will have a p, which is the probability of adopting due to mass media, and a q, which is the probability of adopting due to word-of-mouth effects. In the network version, each agent will also have a set of neighboring agents.

DESIGN CHOICES FOR AGENT-BASED MODELING

1. Scope of the Model – What aspects of the complex system under examination will be described in the model?

2. Agents – What agent types exist in the model?

3. Properties – What properties does each agent have?

4. Behaviors –What behaviors / actions does each agent possess?

5. Environment– What external forces act on each agent, including other agents and the external environment?

6. Input and Output –What inputs to the model exist? What outputs can be collected from the model?

7. Time Step –What is the order of events in the model?

Fig. 3. Design choices for agent-based modeling.

(4) Behaviors — Each agent exhibits a set of behaviors. In our model, each agent makes a decision to adopt on the basis of a probability, $p+q(n_a/n)$, where n_a is the number of innovators who are neighbors and n is the total number of neighbors of the agent.[5]

(5) Environment — This can be a physical environment (e.g., a city), a social environment (e.g., an email network), or a conceptual environment (e.g., a brand space). The environment defines the interaction topology of the agents. In our model, the agent's social neighbors define its environment.

(6) Input and output — It is also necessary to define the inputs to the model, and the output measures that will be gathered. In the original Bass (1969) model, the main inputs are p, q, and m. These parameters are input parameters for the agent-based model as well. In the network version, the network topology (i.e., how individuals are connected) is another input. The output measure of interest is the number of adopters per time step.

(7) Time step — Almost all agent-based models have an initialization step and then an iteration time step, and these two "phases" must be described. In our model, the initialization step creates the agents and the network, if necessary. In each iteration step, each agent determines if it will innovate due to media effects or word-of-mouth or if it will not innovate. If the agent decides to adopt, then the appropriate properties are updated and the process repeats. When all agents have adopted, the model terminates.

Step 3. Construct the model

Model construction is a software engineering process. The goal is to create an implemented version of the model that can be executed computationally and that corresponds with the conceptual model. There are two alternatives to model implementation. The first is to write the model in a general purpose programming language, such as C, Java, or Python. The second is to make use of an extant ABM language/library. There are four widely used ABM toolkits in existence: NetLogo, Repast, Mason, and Swarm. Railsback, Lytinen, and Jackson (2006) provide a detailed review of all four of these platforms. The initial implementation of our model was straightforward based on the design documentation described in Step 2. The model follows the following steps:

1. *Initialization* — The number of agents is created according to the input parameter m. Each agent is given the same p and q values, according to the input parameters. All agents are set to the default state of not having adopted the innovation (i.e., $adopted_i = false$). The agents are then connected to each other on the basis of the specified network (*preferential* vs. *random*) and with the given network density or connection parameter (k).
2. *Adoption decision*[6] — Any agent i who has not yet adopted the innovation (i.e., $adopted_i = false$) decides whether to adopt the innovation. This is done by generating two random variables each time step ($x_{1,i},\ x_{2,i} \sim U(0,1)$) and comparing them to the p and q values, after which the $adopted_i$ state variable is updated appropriately. Specifically, $adopted_i = true$ if:

$$x_{1,i} < p \text{ or } x_{2,i} < q * \frac{n_{a,i}}{n_i}$$

where n_i is the number of neighbors of agent i and $n_{a,I}$ is the number of neighbors of agent i who are in the $adopted = true$ state.
3. *Statistics collection* — The number of agents who have adopted is recorded.
4. *Repeat* — If there are still agents who can adopt the innovation, then go to Step 2; otherwise, terminate the simulation.

To reproduce the results of the Bass (1969) model, the agents were assumed to neighbor all other agents in Step 2. Once these results were implemented and verified, the specific case of examining agents who only observed network neighbors was investigated.

We implemented two standard network structures in the network-based version of our model: (i) a *random* network (Erdos & Renyi, 1959) and (ii) a *preferential attachment*-based (Barabási & Albert, 1999) network.[7] In the *random*

[5] When replicating the original Bass results, each agent is assumed to neighbor every other agent.

[6] Our model differs from the original Bass model in that each agent individually determines if they will adopt the innovation or not. The agents in this model act synchronously; in other words, each agent determines whether to adopt separately from other agents and then updates its status at the end of an iteration after all other agents have had a chance. This differs from the Bass model because it is equivalent to assuming that the number of innovators at time t is dependent on the number of innovators at time $t-1$ and not on the instantaneous number of innovators at time t. This assumption could be relaxed in the future by allowing the model to update asynchronously. The agents could then decide whether to adopt based on their neighbors who have adopted, including any changes in adoption status since the beginning of the time step. This would still not be a continuous model, but it would be closer.

[7] There are many different ways to create scale-free networks because scale-free networks are networks that have a scale-free degree distribution. However, being scale-free says nothing about how the network was created or how the nodes are connected (i.e., assortativity). The mechanism that we use for the networks in this study is the preferential attachment mechanism of Barabási and Albert (1999). All networks generated using preferential attachment must be scale-free, but not all scale-free networks are generated using preferential attachment. Therefore, for the sake of brevity, we will refer to the network as a preferential attachment or preferential network throughout the paper.

network, the parameter of control is the network density, i.e., the percentage of all agents that each agent is connected to. If the network density is 100%, then this network topology is similar to the interaction structure in the original Bass model.[8] In the *preferential attachment* model, the network is based on the Barabási and Albert (1999) model. In this model, some consumers in the network have a disproportionately large number of neighbors, while other consumers have a small number of neighbors. Specifically, the fraction of nodes $P(d)$ in the network that have d connections to other nodes is $P(d) \sim d^{-\gamma}$, where $\gamma = 3$ when using the Barabási–Albert network-generating model.[9] This type of network has been shown to be representative of some real-world social networks (Barabási & Albert, 1999).

The controlling parameter of the preferential model is k, which describes the minimum number of neighbors of any node in the network. If $k = m - 1$, then the model reduces to the original model in which every agent interacts with every other agent.[10]

We implemented the model in NetLogo (Wilensky, 1999). Pseudo-code (a high-level formal description of the model) and complete source code for the model is freely available from the authors upon request.

Step 4. Analyze the model

Once an ABM has been built, verified, and validated, it can be used as an experimental tool. This is discussed in detail in Section 8, but it involves running the model over a range of parameters and analyzing the results. The results can be examined using statistical tests, such as the Pearson correlation, Student's t-test, the Kolmogorov–Smirnov test, or even regressions across the input and output variables. We conducted a number of experiments on the agent-based Bass model, but before we describe those, we will discuss how we ensured that the model had been constructed rigorously.

7. Guidelines for a rigorous model

Two processes fundamentally define the rigor of an agent-based model: verification and validation. It has been said that before a model undergoes verification and validation, it is just a toy; after a model undergoes verification and validation, it is a tool (North & Macal, 2007). Verification is the process by which the implemented model is shown to correspond to the conceptual model, and it is carried out during the design and construction steps of the model development. Validation is the process by which the implemented model is shown to correspond to the real world, and it usually occurs during the model analysis. Both verification and validation provide the possibility of falsifying the proposed model. Models are like hypotheses; a model presents a possible explanation of the way the world works, but the explanation must be tested. Proving a model to be false can be as simple as showing that it rests on an invalid assumption. It is impossible to completely validate or verify a model (Grimm & Railsback, 2005), but this is true for all models and not just for ABM. For instance, though the Bass model of innovation diffusion has been shown to correspond to many real-world scenarios, it is impossible to say it has been fully validated because there are many innovations that it has not been used to model. Validation and verification are carried out to the extent necessary to convince the target audience of the rigor of the model. An agent-based model may not convince many researchers unless they can access the original model and data. Therefore, when using the ABM approach, both the original source code and all data used to validate the model should be made publicly available. Top-quality research journals can facilitate this practice by providing permanent, publicly available repositories for computational models and data.

Verification and validation should be applied to all models, not just agent-based models; however, in the sections below, we will discuss specifically how to carry them out in the context of ABM. Many of these guidelines are applicable to computational modeling in general (e.g., empirical output validation), but several of them are more applicable to models built using the ABM approach (e.g., micro-face validation).

7.1. Verification

Verification determines how well the implemented model corresponds to the conceptual model. There are three important steps in rigorous verification: documentation, programmatic testing, and test cases (Fig. 4). Though verification attempts do not need to be published, developers of a model should record their efforts and provide on request both a record of verification and the original source code so that other researchers can independently verify the model.

Documentation (Fig. 4, Step 1) is the process of creating records that describe both the conceptual model and the implemented model at a level of detail that allows them to be easily compared (North & Macal, 2007). For instance, all of the choices related to a model's design need to be documented. In addition, there should be sufficient documentation within the code of the implemented model that even an unskilled programmer can compare the documentation of the code to the conceptual documentation and understand what parts of the code are responsible for certain parts of the conceptual model (Gilbert, 2008). The documentation in the code should also be sufficient for a trained programmer to ascertain how the code implements the comments. Good documentation can ensure verification by facilitating the comparison between implemented and conceptual models.

Programmatic testing (Fig. 4, Step 2) involves ensuring that the implemented model does what the programmer believes it to do. There are four major ways to carry out programmatic testing: *unit testing*, *code walkthroughs*, *debugging walkthroughs*, and *formal testing*. *Unit testing* (Beck, 2002) is where every unit or section of code has a test written for it. For instance, if the code carries out a complex mathematical equation, there should be a test that makes sure the code generates the appropriate result for a given input. *Code walkthroughs* are when a group of researchers examines the code and the programmer walks the researchers through each step, describing what the code is supposed to do (North & Macal, 2007). The goal is to determine if the logic expresses the concept that the programmer intended. *Debugging walkthroughs* are when the code is run by the programmer and checked at each step to make sure it is generating the correct results (usually using a debugging tool). *Formal testing* uses logic to show that the code must be correct (Dijkstra, 1976). Unfortunately, most agent-based models are sufficiently complicated that using formal testing is difficult (North & Macal, 2007).

Test cases (Fig. 4, Step 3) involve the use of artificially generated data to make sure that the model functions as described. There are four kinds of test cases: *corner cases*, *sampled cases*, *specific scenarios*, and *relative value testing*. *Corner cases* (Gilbert, 2008) are extreme values of the inputs; by examining corner cases, the researcher can make sure the model does not exhibit aberrant behavior in these

[8] It should be noted that it is computationally more efficient to not create a network and just have each agent count the aggregate number of agents that have adopted, rather than actually query each network linkage and determine if that neighbor has adopted the innovation. In fact, that is what we do in the results presented below.

[9] Other scale-free networks generated using alternative mechanisms may have different exponents. In fact, the exponent of the Barabási–Albert model is $2 + 1/k$, where k is the number of connections added by each new node that enters the network, and this is the same k value that we will discuss in this paper.

[10] A fully-connected network, such as the one described, is sometimes called a panmictic model, which literally means "all mixing."

STEPS TO ENSURE RIGOR IN VERIFICATION

1. **Documentation** – Conceptual design and the implemented model should be documented.
2. **Programmatic Testing** – Testing of the code of the model.

    Unit Testing – Each unit of functional code is separately tested.
    Code Walkthroughs – The code is examined in a group setting.
    Debugging Walkthroughs – Execution of the code is stepped through.
    Formal Testing – Proof of verification using formal logic.

3. **Test Cases and Scenarios** – Without using data, model functions are examined to see if they operate according to the conceptual model.

    Corner Cases – Extreme values are examined to make sure the model operates as expected.
    Sampled Cases – A subset of parameter inputs are examined to discover any aberrant behavior.
    Specific Scenarios – Specific inputs for which the outputs are already known.
    Relative Value Testing – Examining the relationship between inputs and outputs.

Fig. 4. Steps to ensure rigor in verification.

situations. For example, when running the model with no agents, or with no interaction between the agents, the results of the model should be easily predictable, and thus the results can be examined independent of empirical data. In a model of product adoption, if the product contains no attributes that appeal to the consumer, then there should be no product adoptions; if there are adoptions, then there is usually a mistake or a bug in the code. *Sampled cases* are a subset of input values chosen in a variety of ways (e.g., randomly, with the design of experiments methods, or with a more complex method that employs an automated exploration tool) (Lambrecht, Ivens, Vandaele, & Miller, 1998). The output of a sampled case is usually not known, but the range of possible outputs is known. Thus, the goal of testing a sampled case is to make sure that the outputs do not exceed these ranges. For instance, the Bass model will eventually result in 100% product adoption if $p>0$. If the model does not produce this result independent of other parameters, then there is a bug in the code. *Specific scenarios* (Gilbert, 2008) are used when there are particular inputs for which the exact output is known as a result of some insight into the conceptual model or previous knowledge from a subject matter expert. For instance, if $q=0$ in the stochastic Bass model, then the adoption rate is simply a product of mass media (p) and the market size (m), and the expected time to full adoption can be determined. If the model deviates too much from this expected time (allowing for stochasticity), then the model is unverified. *Relative value testing* is a method of verification where a relationship is known between an input and an output such that changing an input in one direction will affect the output in a predictable way. For instance, increasing input parameter x decreases output variable y.

In the agent-based Bass model, we carried out all three steps of verification. As for documentation, Verification Step 1 (Fig. 4), the description of the model design within this paper serves as the conceptual model documentation, and the code is also documented. In terms of programmatic testing, Verification Step 2 (Fig. 4), a combination of code walkthroughs, unit testing, and debugging walkthroughs was employed to ensure that the code was behaving as expected. The code consists of eight major routines, each of which was investigated separately to make sure that they work as expected. Some of the functions require other functions as inputs; thus, to test them separately, test values were passed so that the expected behavior from the routine could be predicted and examined. As confidence was established in each function, they were combined together in a step-wise fashion and verified again. By building up the functions in this manner, the growth in the model's complexity is reduced, which facilitates verification. Finally, with regard to test cases, Verification Step 3 (Fig. 4), corner cases were examined by setting $\{p,q\}=\{0.0,1.0\}$ and $\{p,q\}=\{1.0,0.0\}$ for all other input parameter combinations, and by making sure that the agents adopted the innovation as expected.[11] In addition, several (6) sampled cases were randomly chosen to make sure that the model did not exhibit any aberrant behavior, such as a negative adoption rate or a lack of full adoption despite $p>0$. Relative value testing was also employed; p was held constant while q was increased to see if the number of adopters due to word-of-mouth increased relative to the number of adopters due to mass media.[12] The reverse condition was also checked, with q held constant and p increased. Because of these procedures, we can consider the model to be verified to a rigorous degree.

7.2. Validation

Validation is the process of determining how well the implemented model corresponds to reality. Validation has a long history within modeling (Conway, Johnson, & Maxwell, 1959), and specifically within computational modeling (Carley, 1996; Garcia, Rummel, & Hauser, 2007; Knepell & Arangno, 1993). There are four major steps that must occur to consider the model rigorously validated (Fig. 5): micro-face validation, macro-face validation, empirical input validation, and empirical output validation.

Micro-face validation (Fig. 5, Step 1) is the process of making sure that the mechanisms and properties of the model "on face" correspond to real-world mechanisms and properties. For instance, do the individuals in the model correspond in a meaningful way to real-world individuals? Do the actions possessed by a consumer agent correspond to real-world actions? Do consumer agents possess a realistic amount of information?

Macro-face validation (Fig. 5, Step 2) is the process of showing that the aggregate patterns of the model "on face" correspond to real-world patterns (North & Macal, 2007). For instance, do the dynamics of the model correspond to the real world? Does the theory of the model correspond to our current understanding of the real world? In both micro- and macro-face validation, no data are directly compared to the model; instead, the focus is on showing that the general patterns, attributes, and processes have an explainable correspondence to the real world. It is usually sufficient to describe the relationship between the model and the real world to show that it has

[11] In the first case, no adoptions should occur because the only effect is social and no one has adopted in the initial state. In the second case, everyone adopts in the first time step of the model because the probability of adopting is 1.0.

[12] It should be noted that this required the creation of two additional output variables that recorded the number of adopters due to the two processes at each time step. Often such debugging outputs must be created to verify a model.

Please cite this article as: Rand, W., & Rust, R.T., Agent-based modeling in marketing: Guidelines for rigor, *Intern. J. of Research in Marketing* (2011), doi:10.1016/j.ijresmar.2011.04.002

STEPS TO ENSURE RIGOR IN VALIDATION

1. **Micro-Face Validation** – Elements of the implemented model correspond "on face" to the real world.

2. **Macro-Face Validation** – Processes and patterns of the implemented model correspond "on face" to the real world.

3. **Empirical Input Validation** – The data used as inputs to the model corresponds to the real world.

4. **Empirical Output Validation** – The output of the model corresponds to the real world.

   Stylized Facts / Subject Matter Experts – Generally known patterns of behavior that are important to reproduce in the model.
   Real World Data – Recreating real world results using the model.
   Cross-Validation – Comparing the new model to a previous model that has already been validated.

Fig. 5. Steps to ensure rigor in validation.

been validated "on face." An additional level of validation can be gained by having subject matter experts review the model "on face."

Empirical input validation (Fig. 5, Step 3) is the process of ascertaining that the data being input into the model are accurate and bear a correspondence to the real world. In this case, it is sufficient to explain how the data are derived and what correspondence they have to model inputs. For instance, basing the rate of adoption for a new product on the rate of adoption of similar products is an example of empirical input validation. Whenever possible, the inputs of an agent-based model should be calibrated to actual data from the real world. Garcia et al. (2007) provide one method of automating this process for consumer behavior by describing how to validate inputs on the basis of conjoint analysis. Their method makes it possible to create empirically-driven rules that govern agent decisions in the model.

Empirical input validation can extend beyond parameters to the basic behaviors of the model. For instance, Goldenberg et al. (2010) examined how network effects could reduce adoption due to a "chilling effect." Rust (2010) states that this is because the adoption rule that was used contains a chilling effect and suggests an alternative adoption rule. In cases like this, it may be necessary to empirically validate the adoption rule by examining consumer behavior patterns or to test different rules and report the model's sensitivity to this input.

Another way to set the input values is to partition the data into two sets (sometimes called *training* and *test* data) and use one set of data (the *training* data) to *calibrate* the model by manipulating the inputs until the output of the model matches the training set (Grimm & Railsback, 2005). This can be done manually, but a much better way is to automate the process using a machine learning technique, such as a genetic algorithm or neural network. Once the model has been calibrated, the second set of data (the *test* data) can be used as a validation set.

Many other approaches to setting the parameters of a model are possible, such as ethnographic studies (Agar, 2005; Axtell et al., 2002) and behavioral experiments (Castella, Trung, & Boissau, 2005), but a detailed look at all of these methods is beyond the scope of this paper.

Regardless of the method used to calibrate the inputs, a robustness or sensitivity analysis should be conducted to explore the sensitivity of the model to a particular set of inputs. This is done by holding some parameter(s) constant and then varying one or more target parameters to see how sensitive the model results are to each parameter. If the outputs of interest are not dramatically affected by changes in this input, or if they are affected in a predictable way, then the input can be set to a standard value and the rest of the model can be explored. If the model is very sensitive to the input under investigation, then it is necessary to investigate why it is sensitive to that input and see how that affects the result. In the case of the model discussed in this paper, we varied all of the key parameters (m, p, and q) for a set of the experiments (preferential attachment experiments) that were of interest to our research question. This gives the researcher the ability to answer questions like "If m increases, but p and q are left alone, what is the result?" These results are available from the authors upon request.

Empirical output validation (Fig. 5, Step 4) involves showing that the output of the implemented model corresponds to the real world (North & Macal, 2007). This is the key test of a model's validity. Empirical output validation tests the model designer's hypothesis (i.e., the implemented model). There are three ways to empirically validate a model's output: *stylized facts*, *real-world data*, and *cross-validation*. *Stylized facts* are general concepts of a complex system derived from the knowledge of subject matter experts (e.g., all technology adoption curves are s-shaped). If the model's primary purpose is to be a thought experiment, then validating the model against a stylized fact is sufficient to show the validity of the model. *Real-world data validation* is necessary if the model is being used as a predictive model. In this case, there may often be only one real historical data set and many model runs; thus, empirical output validation involves showing that the real world is a possible output of this model (i.e., that the real-world data set lies within the statistical distribution of the model data). If there are many outputs from the model and the real world (e.g., a time series of purchases), then real-world data validation involves showing that the average model results correlate with the real-world results. *Cross-validation* is an optional validation process that compares the new model against another model that has already been validated, even if that other model uses another methodology (e.g., system dynamics or analytical modeling). If the two models produce similar results, then the validity of the new model has been increased.

As described in Section 3, cross-validation is a very important approach for model validation in a field such as marketing, where researchers are more familiar with other techniques such as game theory, statistical regression, and econometric modeling. The ABM approach can be used to recreate the same results as these other methods under similar assumptions, and then those assumptions can be relaxed using ABM in a way that cannot be done using the traditional approach. For instance, in the illustrative example, we will show that under certain assumptions (e.g., that consumers observe the whole population), we can recreate the original Bass (1969) model results and that under additional assumptions, we can produce similar results to a more complex closed form model (Van den Bulte & Joshi, 2007). Once this has been demonstrated, it is more credible to present results that move beyond these simple models, such as the network-based results that we present in this paper.

For our model, we carried out the four steps of validation as described above. At the micro-face level, Validation Step 1 (Fig. 5), the

Table 3
ABM model parameters.

Durable	m	p	q
Refrigerators	40,001	.0026167	.21566
Freezers	21,973	.0181190	.17110
TV	96,717	.0278770	.25105
Softener	5793	.0177030	.29695
AC	16,895	.0103990	.41861
Dryer	15,092	.0172060	.35688
Lawnmowers	44,751	.0091837	.33790
Bed	76,589	.0058760	.24387
Coffee	58,838	.0171350	.30145
Iron	55,696	.0286320	.32791
Player	21,937	.0247960	.65410

Table 4
ABM model correlation with empirical sales data.

Durable	Period of interest	ABM R^2	Bass R^2
AC	1950–1961	0.84	0.90
Bed	1950–1961	0.98	0.93
Coffee	1951–1961	0.75	0.69
Dryer	1950–1961	0.88	0.85
Freezers	1947–1961	0.59	0.47
Lawnmowers	1949–1961	0.96	0.89
Refrigerators	1926–1940	0.62	0.76
TV	1949–1961	0.25	0.07

agents in the model possess a *p* and a *q*, which represent their probability of adopting based on word-of-mouth or mass media; these factors seem “on face” valid. In addition, in the network model, the agents possess a local social network that does not allow them to discover information about the whole population.

At the macro-face level, Validation Step 2 (Fig. 5), the aggregated patterns seem to suggest typical innovation adoption patterns “on face”, with at first a few adopters and then more and more until the whole population has adopted (sigmoid curve). The underlying theory of the model is based on the classic diffusion of innovation literature (Rogers, 1962); thus, to the extent that the literature is valid, the theory of this model is valid.

At the empirical input level, Validation Step 3 (Fig. 5), the input data are drawn from the original Bass (1969) paper. At first glance, parameters *p* and *q* in the agent-based model do not seem to be the same as they are in the original Bass model (Garcia et al., 2007). This is because the original Bass model assumed that *p* and *q* were hazard rates for the whole population, not for individuals. However, if we are interested in examining how the agent-based model behaves in an environment where all individuals can observe the state of all other individuals, then the effects of *p* and *q* in the agent-based model are very similar to their effects in the original Bass model; therefore, we can use the same values to initiate the agent-based model. As for the validity of the networks, the random network is not very representative of real-world networks' structures, but it is more realistic than the models in which every consumer knows every other consumer. The preferential attachment model (Barabási & Albert, 1999) has previously been shown to be representative of some real-world social networks.

The *m* value as originally described by Bass (1969) is the size of the market of potential adopters, which must be known a priori, determined by early market analysis, or estimated. This could correspond directly to the number of agents in the agent-based model, but because decreasing this value only impacts the variance of the results,[13] it is safe to model a smaller number of agents to expedite experimental analysis. This should be investigated for each model to determine whether the number of agents substantially alters the results.

As we alluded to above, there are many different ways to set (i.e., *calibrate*) the parameter values of an agent-based model. In our case, the *p*, *q* and *m* values were drawn directly from the Bass (1969) paper, which in turn was calibrated from earlier data. In the case of the network inputs, we used a different approach; because we were less sure of what inputs to use, we chose to scan a variety of networks that have been reported to be representative of real-world social networks and investigate whether the model was sensitive to different networks.

With regard to empirical output validation, Validation Step 4 (Fig. 5), this model corresponds to certain stylized facts in the diffusion literature. The output always takes on the classic sigmoid shape of innovation adoption when $q>0$. To further validate the model, all 11 cases originally described in the Bass (1969) paper were recreated. The descriptions of the model inputs are listed in Table 3. These values are drawn directly from the original Bass model because in the agent-based model, when the agents can observe the entire population, the model becomes a discrete (both in time and in the granularity of the population) version of the Bass model. As a result, the exact same values of *p*, *q* and *m* are used to allow for comparison with the original Bass model. For each set of inputs, the model was executed 30 times using NetLogo's BehaviorSpace (Wilensky, 2003) tool, and the number of adoptions at each time step was collected in a text file. The results were imported to R (R Development Core Team, 2008) and averaged within each parameter set across all 30 runs. Finally, for forecasting accuracy, the empirical data were imported from the eight datasets still available with sufficient data (http://www.bassbasement.org) for the same time period that Bass (1969) used. Using these data, the r^2 between the empirical data and the model data was calculated. These results, as well as Bass' original results, are illustrated in Table 4. As can be seen, this model achieves results similar[14] to those of the original Bass model, and the ABM r^2 exceeded the Bass results in six out of eight cases, indicating that the model produces a similarly valid description of innovation adoption.

At this point, the model has been sufficiently verified and validated, so it can accurately serve as a model of innovation adoption similar to the original Bass model, and we can now explore the results of the model.

8. Using the model

Once the model has been verified and validated, we are now able to move beyond the model construction to discover new insights from the model. An agent-based model can be used as a computational experiment. To postulate an experiment, the researcher specifies a certain combination of inputs and varies one or more of the inputs to see how the changes affect the outputs. Once a series of experiments has been run, the results from the agent-based model can be analyzed using standard statistical approaches that are applied to any large dataset of empirical data.

One important caveat is that because almost all models constructed using ABM are stochastic in nature, it is important to run the same model multiple times with the same inputs to capture the distribution over possible results that can exist due to the model's stochasticity. Essentially, it is important to observe an event enough times that statistical inference can be made about the relationship between the inputs and the outputs. These experiments are usually carried out using some sort of experimental tool that iterates through all of the inputs specified by the researcher, runs the model repeatedly for each treatment, and then aggregates the results so that they can be analyzed using Excel, R, or another analysis tool.

[13] The variance of these results with the smaller number of agents is still limited enough to allow for the creation of accurate adoption curves.

[14] The differences between the two models could be a result of the ABM version's discrete nature or the fact that the model was implemented as a synchronous model.

Please cite this article as: Rand, W., & Rust, R.T., Agent-based modeling in marketing: Guidelines for rigor, *Intern. J. of Research in Marketing* (2011), doi:10.1016/j.ijresmar.2011.04.002

PREFERENTIAL NETWORK MODEL RESULTS FOR *k=1* TO *k=4*, OVER 30 RUNS. THE DOTTED LINE IS THE ORIGINAL BASS MODEL RESULTS FOR THE SAME PARAMETERS

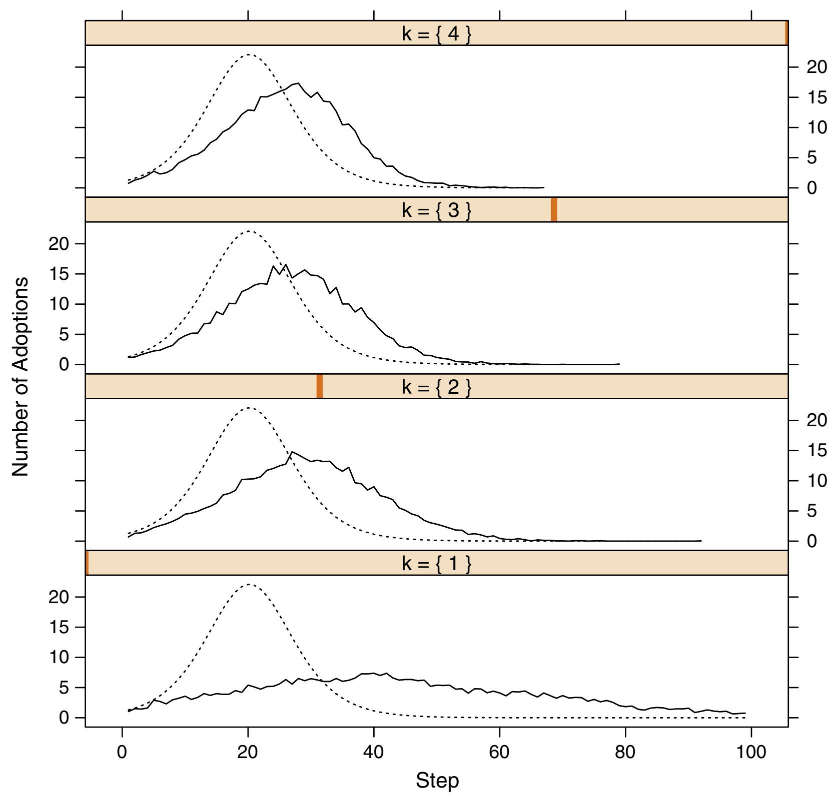

Fig. 6. Preferential network model results for $k=1$ to $k=4$, over 30 runs. The dotted line is the original Bass model results for the same parameters.

For instance, in the innovation adoption model, we will explore how network density and topology affect the adoption of innovations. For all of these experiments, p and q were set the same as the Bass (1969) parameters for a typical durable (refrigerators; $p=0.0026167$ and $q=0.21566$), but m was decreased from 40,001 to 400 because analysis of results is facilitated by allowing faster examination of a large parameter space, and as the variance measures indicate, this had no qualitative influence on the results. In our first experiment, the network parameter was first set to the *random* topology, and the network density parameter was varied from 0% to 100% at 10% increments. For each increment, the model was executed 30 times using NetLogo's headless BehaviorSpace (Wilensky, 2003) mode, and the number of adopters at each time step was collected. The results were imported into R and averaged across runs.

For the random network, the model results are qualitatively similar to the original Bass results, except, of course, when the network density was 0%. A network density of 0% results in agents having no neighbors and so no adoptions occur. For all other network densities, the adoption curve is similar to the Bass model results. This is because the network is relatively dense even at 10%, which results in the average agent having 40 neighbors. Given the random distribution of links, this means the average path length (i.e., average distance between two individuals) across the network is not large (~1.6 individuals), which results in rapid diffusion.

To examine a more realistic network, we use *preferential* topology where the minimal number of neighbors, k, was varied from 1 to 10. For the preferential model, results for $k=1$ to $k=4$ are presented in Fig. 6 because these results are the most starkly different. To test these differences, we utilized the Kolmogorov–Smirnov pairwise test for values of $k \leq 4$.[15]

The preferential attachment mechanism generates hub (highly connected individuals) and spoke (less connected individuals) structures. As a result, at low levels of network connectivity, if a hub adopts the innovation, then it spreads quickly due to word-of-mouth. However, if a hub does not adopt the innovation, then diffusion is driven more by p than q, which in most cases slows the overall adoption rate. As k is increased and m is held constant, it is more likely that any innovation will reach a hub because each increase in k means that all nodes have at least one additional neighbor, and the overall number of neighbors will scale according to a power law.[16]

These results indicate that in networks where consumers do not know many other consumers and relationships are distributed more unequally, innovation diffusion can occur at a much slower rate than the results predicted by the traditional Bass model. However, if the network has a moderately high level of connections, the results of innovation diffusion on a network structure are qualitatively similar to the results of the original Bass model. Prior research with a different model supports this conclusion (Stonedahl, Rand, & Wilensky, 2008).

15 Due to convergence, we do not present any results for $k>4$.

16 At first glance, this result may appear contrary to Watts and Dodds (2007), but their results are relative to the size of the hub (i.e., do influentials create proportionally more adoption given the number of links that they have) and not the absolute rate of adoption. Moreover, for most of their results they use a deterministic threshold functional form (Watts, 2002) and not a probability distribution for the adoption decision, which makes the results difficult to compare directly.

Please cite this article as: Rand, W., & Rust, R.T., Agent-based modeling in marketing: Guidelines for rigor, *Intern. J. of Research in Marketing* (2011), doi:10.1016/j.ijresmar.2011.04.002

Table 5
Statistical regression of ρ^* on m, p, q, and k.

	Estimate	Std. error	t-value
Intercept	−80.54	1.896	−42.47***
m	.08389	.0003573	234.81***
p	11.03	.1799	61.30***
q	669.2	23.55	28.41***
k	106.7	3.925	27.18***

*** p<0.01.

Though the results above are sufficient to answer the question as to whether social network structures can affect diffusion curves with a positive answer, the researcher may want to explore the relationship between a larger variety of inputs and outputs of an agent-based model. For instance, the results of an agent-based model can be viewed as a large panel dataset and econometric techniques can be applied. For instance, one example would be to run a regression across the inputs and the outputs, treating the inputs as independent variables and the outputs as dependent variables. To briefly illustrate this possibility, we created a new experiment in which we used only the preferential attachment network and varied the p (6 levels), q (6 levels), m (4 levels) and k (4 levels) inputs. We also ran each treatment 10 times for a total of 5760 different elements of our dataset. We then regressed the peak adoption rate (ρ^*) in each of these setups across the input parameters of p, q, m and k using a generalized linear model similar to the method used by Stephen and Berger (2009) and Stephen et al. (2010):

$$\rho^* = \beta_0 + \beta_1 p + \beta_1 q + \beta_3 m + \beta_4 k.$$

The results are presented in Table 5. As is clear from the results for this set of inputs, all of the parameters have a statistically significant effect on the peak adoption rate. This confirms our findings that density of the preferential attachment network, as controlled by parameter k, has a significant effect on the adoption curve and, in this case, specifically on the peak adoption rate. This is just one example of how traditional econometric approaches can be applied to better understand computational methods. Further investigations into this combination of methods are definitely warranted in future work.

9. Conclusion

We have presented guidelines for the rigorous use of agent-based modeling (ABM) in marketing along with an illustrative example of how to use ABM to understand the diffusion of innovations. Despite the difficulties of building rigorous agent-based models, the benefits of a modeling approach for marketing that focuses on the individual level and observes the aggregate results are substantial. ABM provides an approach that enables the concretization of many theories of consumer behavior in a way that is measurable and testable, while at the same time permits the inclusion and comparison of empirical data. Moreover, the realism of ABM facilitates model comprehension by managers and stakeholders. As the cost of computational power continues to decline and the ability to create more and more elaborate models continues to increase, ABM and other computationally intensive methods are likely to assume a larger role in our understanding of the world. Therefore, it is important for marketing researchers to understand how to rigorously use this new method.

Acknowledgments

Support for Rand was partially provided by National Science Foundation IIS Grant #1018361. We also thank the Center for Complexity in Business at the Robert H. Smith School of Business at the University of Maryland for its support.

References

Agar, M. (2005). Agents in living color: Towards emic agent-based models. *Journal of Artificial Societies and Social Simulation*, *8*(1), 4.
Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Review*, *103*(3), 411–423.
Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *American Economic Review (Papers and Proceedings)*, *84*, 406–411.
Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.
Axelrod, R. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration.* Princeton, NJ: Princeton University Press.
Axtell, R. L., Epstein, J. M., Dean, J. S., Gumerman, G. J., Swedlundx, A. C., Harburger, J., et al. (2002). Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proceedings of the National Academy of Sciences*, *99* (Suppl. 3), 7275–7279.
Barabási, A. -L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*, 509–512.
Barr, J., & Saraceno, F. (2002). A computational theory of the firm. *Journal of Behavior and Organization*, *49*(3), 345–361.
Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science*, *36*(9), 1057–1079.
Batty, M. (2005). *Cities and complexity: Understanding cities with cellular automata, agent-based models and fractals.* Cambridge, Massachusetts: MIT Press.
Beck, K. (2002). *Test driven development: By example.* : Addison-Wesley Professional.
Berger, T. (2001). Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics*, *25*(2), 245–260.
Brown, D. G., Riolo, R., Robinson, D. T., North, M., & Rand, W. (2005). Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical Systems*, *7*(1), 25–47 Special Issue on Space–Time Information Systems.
Carley, K. M. (1996). *Validating computational models.* Working Paper. Carnegie Mellon University.
Castella, J. -C., Trung, T. N., & Boissau, S. (2005). Participatory simulation of land-use changes in the northern mountains of Vietnam: The combined use of an agent-based model, a role-playing game and a geographic information system. *Ecology and Society*, *10*(1), 27.
Casti, J. L. (1995). Seeing the light at El Farol. *Complexity*, *5*(1), 7–10.
Churchill, G. A., Jr. (1979, Februaryy). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, *XVI*, 64–73.
Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, *17*, 1–25.
Conway, R. W., Johnson, B. M., & Maxwell, W. L. (1959 October). Some problems of digital systems simulation. *Management Science*, *6*(1), 104–105.
Deffuant, G., Huet, S., & Amblard, F. (2005). An individual-based model of innovation diffusion mixing social value and individual benefit. *The American Journal of Sociology*, *110*(4), 1041–1069.
Delre, S. A., Jager, W., Bijmolt, T. H. A., & Janssen, M. A. (2007, August). Targeting and timing promotional activities: An agent-based model for the takeoff of new products. Special Issue: Complexities in Markets. *Journal of Business Research*, *60*(8), 826–835.
Delre, S. A., Jager, W., Bijmolt, T. H. A., & Janssen, M. A. (2010, March). Will it spread or not? The effects of social influences and network topology on innovation diffusion. *Journal of Product Innovation Management*, *27*(2), 267–282.
Dijkstra, E. (1976). *A discipline of programming.* Englewood Cliffs, NJ: Prentice-Hall.
Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up.* Boston: MIT Press.
Erdos, P., & Renyi, A. (1959). On random graphs. *Publications Math Debrecen*(6), 290–297.
Frels, J. K., Heisler, D., Reggia, J. A., & Schuetze, H. -J. (2006, June). Modeling the impact of customer interactions in technology markets. *Journal of Cellular Automata*, 91–103.
Garber, T., Goldenberg, J., Libai, B., & Muller, E. (2004). From density to destiny: Using spatial dimension of sales data for early prediction of new product success. *Marketing Science*, *23*(3), 419–428.
Garcia, R. (2005). Uses of agent-based modeling in innovation/new product development research. *Journal of Product Innovation Management*, *22*(5), 380–398.
Garcia, R., & Jager, W. (2011). Agent-based modeling of innovation diffusion. *Journal of Product Innovation Management*, *28*(2), 148–151.
Garcia, R., Rummel, P., & Hauser, J. (2007, August). Validating agent-based marketing models using conjoint analysis. Special Issue: Complexities in Markets. *Journal of Business Research*, *60*(8), 848–857.
Gerbing, D. W., & Anderson, J. C. (1988, May). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, *XXV*, 186–192.
Gilbert, N. (2008). *Agent-based models.* Sage Publications.
Gilbert, N., Jager, W., Geffuant, G., & Adjali, I. (2007). Complexities in markets: Introduction to the special issue. *Journal of Business Research*, *60*(8), 813–815.
Gilbert, N., & Troitzsch, K. (2005). *Simulation for the social scientist: Second edition.* Open University Press.
Goldenberg, J., Han, S., Lehmann, D. R., & Hong, J. W. (2009). The role of hubs in the adoption process. *Journal of Marketing*, *73*(2), 1–13.
Goldenberg, J., Libai, B., Moldovan, S., & Muller, E. (2007). The NPV of bad news. *International Journal of Research in Marketing*, *24*, 186–200.
Goldenberg, J., Libai, B., & Muller, E. (2001a). Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new

product growth through stochastic cellular automata. *Academy of Marketing Science Review*, *9* special issue on Emergent and Co-Evolutionary Processes in Marketing.
Goldenberg, J., Libai, B., & Muller, E. (2001b). Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, *12*(3), 209–221.
Goldenberg, J., Libai, B., & Muller, E. (2002, Aprill). Riding the saddle, how cross-market communications creates a major slump in sales. *Journal of Marketing*, *66*, 1–16.
Goldenberg, J., Libai, B., & Muller, E. (2010). The chilling effect of network externalities. *International Journal of Research in Marketing*, *27*(1), 4–15.
Goldenberg, J., Libai, B., Solomon, S., Jan, N., & Stauffer, S. (2000). Marketing percolation. *Physica A*, *284*(1–4), 335–347.
Grimm, Volker, Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*, 115–126.
Grimm, V., & Railsback, S. F. (2005). *Individual-based modeling and ecology*. Princeton: Princeton University Press.
Heppenstall, A., Evans, A., & Birkin, M. (2006). Using hybrid agent-based systems to model spatially-influenced retail markets. *Journal of Artificial Societies and Social Simulation*, *9*(3). http://jasss.soc.surrey.ac.uk/9/3/2.html
Hill, R., & Watkins, A. (2007, Septemberr). A simulation of moral behavior within marketing exchange relationships. *Journal of the Academy of Marketing Science*, *35*, 417–429.
Hill, R., & Watkins, A. (2009). The profit implications of altruistic versus egoistic orientations for business-to-business exchanges. *International Journal of Research in Marketing*, *26*(1), 52–59.
Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. Reading, MA: Addison-Wesley.
Irwin, E. G., & Geoghegan, J. (2001). Theory, data, methods: Developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and Environment*, *85* (1–3), 7–24.
Knepell, P. L., & Arangno, D. C. (1993). *Simulation validation: A confidence assessment methodology.* IEEE Computer Society Press.
Lambrecht, M. R., Ivens, P. L., Vandaele, N. J., & Miller, J. H. (1998). Active nonlinear tests (ants) of complex simulation models. *Management Science*, *44*(6), 820–830.
LeBaron, B. (2000). Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, *24*(5–7), 679–702.
Libai, B., Muller, E., & Peres, R. (2005). The role of seeding in multi-market entry. *International Journal of Research in Marketing*, *22*, 375–393.
Libai, B., Muller, E., & Peres, R. (2009). The role of within-brand and cross-brand communications in competitive growth. *Journal of Marketing*, *73*(3), 19–34.
Lusch, R. F., & Tay, N. (2004). Agent-based modeling: Gaining insight into firm and industry performance. In Moorman Christine, & Donald R. Lehman (Eds.), *Assessing marketing strategy performance* (pp. 213–227). Cambridge, MA: Marketing Science Institute.
Marks, R. E., Midgley, D. F., & Cooper, L. G. (1997). Breeding competitive strategies. *Management Science*, *43*(3), 257–275.
Marks, R. E., Midgley, D. F., & Cooper, L. G. (2006). Co-evolving better strategies in oligopolistic price wars. In Rennard Jean-Philippe (Ed.), *Handbook of research on nature-inspired computing for economy and management*. Hershey, PA: Idea Group Inc. Ch. 52.
Marks, R. E., Midgley, D. F., Cooper, L. G., & Shiraz, G. M. (1999). The complexity of competitive marketing strategies. *Complexity International*, *6*.
Miller, J. H., & Page, S. (2007). *Complex adaptive systems: An introduction to computational models of social life.* Princeton University Press.
Moldovan, S., & Goldenberg, J. (2003). Cellular automata modeling of resistance to innovations: Effects and solutions. *Technological Forecasting and Social Change*, *71*(5), 425–442.
North, M. J., & Macal, C. M. (2007). *Managing business complexity: Discovering strategic solutions with agent-based modeling and simulation.* Oxford University Press.
North, M. J., Macal, C. M., St. Aubin, J., Thimmapuram, P., Bragen, M., Hahn, J., et al. (2010). Multi-scale agent-based consumer market modeling. *Complexity*, *15*(5), 37–47.
Parker, D. C., Brown, D. G., Polhill, J. G., Deadman, P. J., & Manson, S. M. (2008). Illustrating a new 'conceptual design pattern' for agent-based models and land use via five case studies: The MR POTATOHEAD framework. In A. Lopez Paredes, & C. Hernandez Iglesias (Eds.), *Agent-based modelling in natural resource management* (pp. 23–51). Valladolid, Spain: Universidad de Valladolid.
Parunak, H. V. D., Savit, R., & Riolo, R. L. (1998). Agent-based modeling vs. equation-based modeling: A case study and user's guide. In J. Sichman, R. Conte, & N. Gilbert (Eds.), *Multi-agent systems and agent-based simulation (lecture notes in computer science)* (pp. 10–25). Springer.
Polhill, G. J., Parker, D., Brown, D., & Grimm, V. (2008). Using the ODD protocol for describing three agent-based social simulation models of land-use change. *Journal of Artificial Societies and Social Simulation*, *11*(2). http://jasss.soc.surrey.ac.uk/11/2/3.html
Prietula, M., Carley, K. M., & Gasser, L. (1998). *Simulating organizations: Computational models of institutions and groups.* Cambridge, MA: The MIT Press.
R Development Core Team (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing3-900051-07-0. http://www.R-project.org
Rahmandad, H., & Sterman, J. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science*, *54*(5), 998–1014.
Railsback, S. F., Lytinen, S. L., & Jackson, S. K. (2006). Agent-based simulation platforms: Review and development recommendations. *Simulation*, *82*(9), 609–623.
Rand, W. (2006). *Machine learning meets agent-based modeling: When not to go to a bar. Agent 2006, Chicago, IL, USA.*
Rogers, E. M. (1962). *Diffusion of innovations.* Free Press.
Rust, R. (2010). Network externalities—Not cool? A comment on 'The chilling effects of network externalities'. *International Journal of Research in Marketing*, *27*(1), 18–19.
Schelhorn, T., O'Sullivan, D., Haklay, M., & Thurstain-Goodwin, M. (1999). STREETS: An agent-based pedestrian model. *CASA working papers 9. Center for Advanced Spatial Analysis*. London, UK: University College London.
Shaikh, N. I., Ragaswamy, A., & Balakrishnan, A. (2005). *Modelling the diffusion of innovations using small world networks.* Working Paper.
Sharara, H., Rand, W., & Getoor, L. (2011, July 17–21). Differential adaptive diffusion: Understanding diversity and learning whom to trust in viral marketing. *Proceedings of the AAAI International Conference on Weblogs and Social Media (ICWSM)* Barcelona, Spain.
Siebel, F., & Kellam, L. (2003). The virtual world of agent-based modeling: Procter and Gamble's dynamic supply chain. *Perspectives on Business Innovation*, *9*, 22–27.
Stephen, A. T., & Berger, J. A. (2009). Creating contagious: How social networks and item characteristics combine to drive persistent social epidemics (April 9). http://ssrn.com/abstract=1354803 Available at SSRN
Stephen, A. T., Dover, Y., & Goldenberg, J. (2010, May 17). *A comparison of the effects of transmitter activity and connectivity on the diffusion of information over online social networks*. INSEAD Working Paper No. 2010/35/MKT. http://ssrn.com/abstract=1609611 Available at SSRN
Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world.* Irwin/McGraw-Hill.
Stonedahl, F., Rand, W., & Wilensky, U. (2008). Multi-Agent learning with a distributed genetic algorithm. *Proceedings of Autonomous Agents and Multi-Agents Systems (AAMAS) ALAMAS + ALAg Workshop, May 12–16, 2008, Estoril, Portugal*.
Tay, N. S. P., & Lusch, R. F. (2002). Agent-based modeling of ambidextrous organizations: Virtualizing competitive strategy. *IEEE Transactions on Intelligent Systems*, *22*(5), 50–57.
Tay, N. S. P., & Lusch, R. F. (2005). A preliminary test of Hunt's General Theory of Competition: Using artificial adaptive agents to study complex and ill-defined environments. *Journal of Business Research*, *58*(9 SPEC. ISS), 1155–1168.
ThinkVine (2010). How we use agent-based modeling in our marketing simulation & planning tool. Technical white paper. http://www.thinkvine.com/
Torrens, P. M. (in press). Moving agent-pedestrians through space and time. *Annals of the Association of American Geographers*.
Toubia, O., Goldenberg, J., & Garcia, R. (2008). A new approach to modeling the adoption of new products: Aggregated diffusion models. *MSI Reports: Working Papers Series, No. 08-001* (pp. 65–76).
Valente, T. W. (1995). *Network models of the diffusion of innovations.* Cresskill, NJ, USA: Hampton Press, Inc.
Van den Bulte, C., & Joshi, Y. V. (2007). New product diffusion with influentials and imitators. *Marketing Science*, *26*(3), 400–421.
Walsh, W., & Wellman, M. (1999). Modeling supply chains formation in multiagent systems. *Proc. IJCAI-99 Workshop Agent-Mediated Electronic Commerce*.
Watkins, A., & Hill, R. (2009). A simulation of business-to-business decision making in a relationship marketing context. *Industrial Marketing Management*, *28*(8), 994–1005.
Watts, D. J. (1999). *Small worlds: The dynamics of networks between order and randomness.* Princeton: Princeton University Press.
Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceeding of the National Academy of Sciences*, *99*, 5766–5771.
Watts, D. J., & Dodds, P. S. (2007). Influentials, networks and public opinion formation. *Journal of Consumer Research*, *34*(4), 441–458.
Wilensky, U. (1999). *Center for connected learning and computer-based modeling.* Evanston, IL: Northwestern University. http://ccl.northwestern.edu/netlogo
Wilensky, U. (2003). *BehaviorSpace.* Evanston, IL: Center for Connected Learning and Computer Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/behaviorspace
Wilkinson, I., & Young, L. (2002). On cooperating: Firms, relations, networks. *Journal of Business Research*, *55*, 123–132.

Please cite this article as: Rand, W., & Rust, R.T., Agent-based modeling in marketing: Guidelines for rigor, *Intern. J. of Research in Marketing* (2011), doi:10.1016/j.ijresmar.2011.04.002