

Chapter 9

Exploring Interventions on Social Outcomes with In Silico, Agent-Based Experiments



Flaminio Squazzoni and Federico Bianchi

Abstract Agent-Based Modeling (ABM) is a computational method used to examine social outcomes emerging from interaction between heterogeneous agents by computer simulation. It can be used to understand the effect of initial conditions on complex outcomes by exploring fine-grained (multiple-scale, spatial/temporal) observations on the aggregate consequences of agent interaction. By performing *in silico* experimental tests on policy interventions where *ex ante* predictions of outcomes are difficult, it can also reduce costs, explore assumptions and boundary conditions, as well as overcome ethical constraints associated with the use of randomized controlled trials in behavioral policy. Here, we introduce the essential elements of ABM and present two simple examples where we assess the hypothetical impact of certain policy interventions while considering different possible reactions of individuals involved in the context. Although highly abstract, these examples suggest that ABM can be either a complement or an alternative to behavioral policy methods, especially when understanding social processes and exploring direct and indirect effects of interventions are important. Prospects and critical problems of these *in silico* policy experiments are then discussed.

Learning Objectives

By studying this chapter, you will:

- Learn the basic concepts and methodological principles of agent-based modeling.
- Understand the advantages of agent-based modeling compared to other research methods when examining social dynamics.
- Understand how to design agent-based modeling for *in silico* experiments.
- Understand the importance of agent-based modeling for policy appraisal.
- Practice with two examples of agent-based modeling for policy experiments.

F. Squazzoni (✉) · F. Bianchi

Department of Social and Political Sciences, University of Milan,
Milan, Italy, Via Conservatorio 7, 20122

e-mail: flaminio.squazzoni@unimi.it; federico.bianchi1@unimi.it

© The Author(s) 2023

A. Damonte, F. Negri (eds.), *Causality in Policy Studies*, Texts in Quantitative Political Analysis, https://doi.org/10.1007/978-3-031-12982-7_9

217

9.1 Introduction

Behavioral science methodology, including randomized controlled trials (RCTs), is increasingly being used in public policy as a gold standard to estimate causal relationships between interventions and outcomes (e.g., Shafir, 2012; Str a bheim & Beck, 2019). Examples of behavioral policies, from public health to education, have shown the malleability of individual preferences and decisions, as well as the sensitivity of targeted individuals to cognitive frames in responding to policy interventions (Galizzi & Wiesen, 2018). The profound non-linear relationships between policy stimuli and observable and measurable people’s responses, which impinge the mantra of ‘big stimuli vs. big outcomes’ of conventional policy (Squazzoni, 2014), has suggested that if well-conjectured and ‘incentive compatible’, even minimal interventions could cause large-scale outcomes (Dolan & Galizzi, 2014).

The reason why RCTs are considered the “gold standard” in behavioral policy is that random assignment of a representative, targeted population to control and treatment groups, differing only in their manipulated conditions and the identification of any controllable, salient confounding factors by *ex ante* design, are instrumental to estimate causal effects. However, besides fundamental criticism on the often neglected influence of implicit assumptions on unobservable processes in research design (e.g., Imai et al., 2008), the use of experimental methods for public policy has also important pragmatic limitations.

On the one hand, whenever feasible, RCTs for public policy purposes could have a negative benefit-cost ratio. Indeed, ethical obstacles can prevent group selection or the exploration of conditions that would introduce inequality and negative externalities for certain groups. Secondly, economic costs are often severe even for small-scale pilots. Furthermore, the intrusive, ‘outside-in’ nature of experimental policies can affect real-life outcomes and people’s behavior in other domains beyond any intended purpose. This is indeed a fundamental problem: not only do people often react unpredictably and adaptively to interventions (note that this has been a key argument for supporters of behavioral policies against the traditional policy framework based on positive/negative incentives and ‘rational’ response), individuals are also embedded in social contexts so that their exposure to policy treatments can trigger positive and negative network externalities or knowledge spillovers, which might also affect outcome measurements (Dolan & Galizzi, 2015; Squazzoni, 2017). Disentangling any established causal effect between interventions and outcomes in such situations is difficult.

Finally, as suggested by Battistin & Bertoni in Chap. 3, inferences on causal effects of policy interventions would require counterfactual procedures to assess what would have happened to the estimated outcomes had these interventions not taken place. Besides the difficulty of isolating a control group in social reality and introducing a placebo-like neutral information in behavioral policies, endogenous social forces and processes cannot be suspended during a policy experiment. Treating data in a quasi-experimental way by randomization, instrumental variation and discontinuity design can increase the robustness of estimates, thus improving

the internal and external validity of causal inferences. Here, we suggest a complementary strategy: the use of agent-based modeling (ABM) as *in silico* experiments accompanying, augmenting, or even substituting RCTs—whenever needed—in the traditional toolbox of the experimentalist policy analyst.

This policy function of ABM is key especially when: (a) there are no or insufficient empirical data on which to corroborate estimated causal relationships and perform *ex post*, counterfactual assessments; (b) the economic, social, or political costs of RCTs for policy appraisal or assessment are hardly sustainable; (c) ‘social experimenters’ are interested not only in estimating outcomes but also understanding generative processes; (d) there is added value in exploring extreme, boundary, or counterfactual conditions that either do not exist in reality or have not yet occurred but in principle could. In all these cases, we argue that ABM is the only alternative to *ex post* observational analysis to explore and quantify hypothesized relationships between policy interventions and social outcomes. What is lost in terms of empirical realism is gained in terms of understanding the possible generative processes.

Reviews on recent applications of ABMs in various fields, from public health (Giabbanelli et al., 2021; Tracy et al., 2018) to agriculture (Kremmydas et al., 2018) and energy consumption (Klein et al., 2019), have shown that ABM is particularly suitable for providing insights into causal mechanisms, potentially linking interventions to outcomes. By generating “artificial data” via computer simulation, models can help to: (a) explore cases of multiple realizability (i.e., the same effect generated by different social causes and paths), (b) build ‘what-if’ scenario analysis that supports inferences about interventions-outcomes without impacting the targeted population; (c) estimate ‘interference’, network effects and spillovers of policy interventions (e.g., the situation in which one individual’s exposure affects other individuals’ outcomes); and (d) measure possibly multiple direct and indirect outcomes of the same intervention (Chalabi & Lorenc, 2013; Murray et al., 2021; Powell et al., 2017).

While most research has outlined the differences between ABM and more conventional policy approaches and methods, e.g., RCTs (e.g., Gilbert et al., 2018), here we would like to discuss complementarities and potential synergies between various experimental approaches. Indeed, as exemplified by Bravo et al. (2012), by using the computer as an ‘artificial experimental environment’, model parameters can be calibrated on existing individual (experimental) data to perform *in silico* counterfactual tests on any established causal relationship by quantifying the effect of varying initial conditions, especially those that could not be estimated empirically. What could happen to the observed causal relationship between *A* (intervention) and *B* (outcome), if certain hypothesized conditions *C* (either observable or not) were different? Why would *A* necessarily lead to *B* given that *C* may include adaptive, unpredictable individual behavior? As suggested by Manzo (2022), this is not only a problem of internal vs. external validity of estimated relationships (the effect of *A* on *B* would be contingent to a specific empirical instance with all due problems of generalization). It implies a search for causal or dependence relationships of interest not only within data but also via formalized models of “generative mechanisms” that consider mediating behavior and processes on which we might

not have any data. Why and how, when exposed to *A* and under interaction effects that typically occur in social contexts, would individuals behave in such a way to ‘cause’ the emergence of *B*?

The rest of the chapter is organized as follows: In Sect. 9.2, we provide a brief introduction to ABM, by highlighting their specificity compared to other modeling approaches. In Sect. 9.3, we present some hypothetical policy cases on which the advantages of ABM can be understood. Model code is provided to help the reader to understand the potential of ABM for: (1) exploring the effect of parameter variations on the emergence of social outcomes; (2) building alternative scenarios to understand the effect of individual reactions on social outcomes. In Sect. 9.4, we summarize the main contributions of the chapter and discuss critical points and possible developments. Indeed, besides the (many) positive aspects, ABM has also certain weaknesses, including problems of model resolution, empirical validation, and external validity, which all require careful scrutiny.

9.2 Agent-Based Modeling

Agent-based modeling is a “computational method that enables a researcher to create, analyze, and experiment with models composed of agents that interact within an environment” (Gilbert, 2008). Agents may represent individuals, households, organizations, or any other entities, whose actions depend on conditional or stochastic decision-making rules (Bianchi & Squazzoni, 2015; de Marchi & Page, 2014; Macy & Willer, 2002; Tesfatsion & Judd, 2006). Agents can adapt their behavior in response to their own experience (e.g., learning), the interaction with other agents or in response to changes in the environment—e.g., policy interventions (Gilbert & Troitzsch, 2005; Squazzoni, 2012; Tracy et al., 2018).

As dynamic and process-based, ABMs are ideal to study the effects of complex interactions between micro- and macro-levels by exploring ‘generative explanations’ of social outcomes (Epstein, 2006; Hedström & Bearman, 2009; Macy & Flache, 2009). This is especially important in the case of complex adaptive social systems, whose stochastic, non-linear behavior can seldom be mathematically tractable and cannot be estimated deductively without computer simulation exploring various initial conditions and possible input/output paths (Miller & Page, 2009).

Unlike statistical models, which concentrate on relations between aggregate factors (Bianchi & Squazzoni, 2020), ABM starts from representing individual behavior and ends up exploring aggregate dynamics from agent interaction via computer simulation. Social regularities and patterns are neither derived by estimating the values of stochastic parameters that would maximize a model’s fitness to observed data, nor obtained by assumptions on aggregate properties that do not consider individual-level differences (e.g., Hedström & Manzo, 2015; Hedström & Udehn, 2009). ABM parameters are not estimated a posteriori, they are manipulated a priori following an experimental rather than an observational research design (Squazzoni, 2012).

Indeed, instead of being inferred from (or tested against) empirical data, the model allows us to explore hypothesized micro-social processes according to this Coleman-like connection: (a) initial macro parameter conditions → (b) heterogeneous individual behavior → (c) interaction effects → (d) social outcomes (Coleman, 1990). In line with the so-called ‘analytical sociology’ agenda (Hedström & Bearman, 2009; Hedström & Manzo, 2015; Manzo, 2022), ABMs can be viewed as generative models ensuring a high degree of internal validity regarding the “generative sufficient conditions” leading from (a) to (d) via the manipulation of (b) and (c) (Epstein, 2006). Unlike statistical models, generative explanations via ABM does not require the independence of observations as they aim to explore systemic, interdependent social processes, i.e., specific configurations of (a), (b), and (c) that would determine (d). Furthermore, ABM allows us to explore various patterns of agent interaction directly within explicitly represented network structures (Macy & Flache, 2009).

While traditional equation-based models condense either a ‘representative’, collective agent or a homogenous population into stochastic parameters (e.g., think about the modeling tradition in either standard economics or demography), ABM explicitly considers a population of heterogeneous, autonomous agents with different features and decision-making rules who interact either directly or indirectly while being exposed to various environmental stimuli, typically manipulated by the model maker (Gilbert, 2008; Macy & Flache, 2009; Macy & Willer, 2002; Squazzoni, 2012). By running experiments with human subjects, experimentalists aim to test theoretically deduced hypotheses on cause–effect relationships by manipulating the occurrence of an *explanans* (i.e., the treatment) in a randomized sample of individuals and studying the control vs. treatment group differences in the *explanandum*. In a similar fashion, an experimenter can use ABM to run several instances of a model by manipulating the *explanans*—i.e., changing the related model parameters—and then studying any differences in the simulated outcome. Instances could be designed as ‘group-treatment’ policy correlates, artificial agents (whose behavior could be empirically inferred from experimental data, if the ABM exercise is combined with a behavioral experiment, or theoretically postulated if data is not available) would be the correlates of experimental subjects, and their group-level reactions would be the outcome measurement. As such, the computer is used as an artificial laboratory where theoretically derived hypotheses are tested in silico by comparing a baseline (control group) initialization with manipulated scenarios (treatments) where the only difference is the introduction of a possible *explanans* (Squazzoni, 2012).

However, this does not constrain ABM to ‘thought experiments’ (Axelrod, 1997). Quantitative (e.g., population size, resources, network positions) and qualitative parameters (e.g., rules of behavior) related to (a), (b), and (c) can be calibrated according to empirical data (i.e., *empirical calibration*), and aggregate artificial outcomes (d) can be compared to empirical time series or distributions to adjudicate among potential configurations of (a), (b), and (c) those with higher explanatory power (i.e., *empirical validation*) (Boero & Squazzoni, 2005).

9.3 Exploring Artificial Policy Scenarios

In this section, we provide some abstract examples from our own research to illustrate the ABM approach to policy scenarios. Although there are many examples of concrete applications of ABM for policy interventions or design (e.g., Gilbert et al., 2018), here we have summarized two recent contributions that describe our idea of *in silico* experiments.

9.3.1 *Interventions to Increase Competition or Collaboration in Science*

Today, academic life is characterized by a “publish or perish” ethos and growing competition for funds and academic career (Edwards & Siddhartha, 2017; Grimes et al., 2018). While competition is expected to stimulate the quality of publications, scientists must also collaborate especially in reviewing manuscripts before publication to defend robust academic standards of knowledge. This is the important function of “peer review”: vetting scientific manuscripts submitted by authors for publication to a journal by voluntary collaboration of experts guided by journal editors. Unfortunately, research has shown that lack of material incentives or a weak system of symbolic rewards can undermine peer review, as scientists would reduce time and effort in reviewing (typically voluntary and not rewarded), to maximize their efforts in new publishable research which funds, prestige, and career depend on.

Suppose that you are a policymaker wanting to test certain possible interventions to increase cooperation among scientists, but who also want to ensure that this does not compromise the quality of publication. Here are two examples of possible research policy interventions. The first represents a policymaker wanting to increase quality signals of publication so to induce scientists to compete for excellence, e.g., promoting only those scientists who publish in top journals. The second wants to reward peer reviewing by introducing an open science policy that would induce journals to shift from confidential to open peer review so that the identity of any reviewer is public, regardless of the final decisions on manuscripts. This would permit reviewers to claim their review as a reward. Note that even if abstract, both policy interventions are ‘realistic’: scientists are increasingly exposed to competitive rewards under the dominant rhetoric of excellence and comprehensive evaluation in almost all institutional contexts (e.g., Forsberg et al., 2022). In the second case, scientific associations and certain publishers have started to introduce open peer review policies as a means to recognize and reward reviewers (Bravo et al., 2019). Therefore, these examples are abstract (i.e., there is no ‘real policy maker’ commissioning a computational test of such policies) but not completely unrealistic (i.e., these interventions have been explored more locally and by trial and error).

Suppose we prepare a model to test these possible interventions. Assume a population of n agents representing a community of scientists. Assume that scientists are

hired by academic organizations that periodically provide them with some minimal funding R_i (e.g., laboratory equipment, access to online resources, etc.), allocated from a fixed overall amount of resources, $R = \sum_i R_i$. Assume that scientists are required to publish manuscripts to get more funds, reputation, prestige, and career, but that journals are competitive and so accept only a fixed proportion (P) of submitted manuscripts depending on a quality ranking determined by reviewers. Scientists then update their resource share according to their publication record as follows:

$$R_i = \frac{p_i}{\sum_j p_j} R$$

Suppose that, at each time step (t), scientists are required to perform two tasks, i.e., submitting their manuscripts to journals and reviewing manuscripts submitted by others (for the sake of simplicity, let us assume that each manuscript is submitted by only one author and is reviewed by only one reviewer; for a similar model, where we varied the number of reviewers, see Bianchi & Squazzoni, 2016). Assume that time is a scarce resource and both tasks are costly in that scientists need to decide how to allocate their resources between these two tasks.

Assume that the quality of submitted manuscripts (Q_i^s) and review reports (Q_i^r) linearly depends on the amount of resources allocated by scientists to these two tasks, as in:

$$Q_i^s = e_i R_i$$

$$Q_i^r = R_i - Q_i^s = (1 - e_i) R_i,$$

where e_i determines how resources are allocated between submitting and reviewing.

Following Squazzoni and Gandelli (2012, 2013), we assume that reviews may be biased, so the actual quality of manuscripts could be only approximated by the reviewer depending on the level of resources individually invested by the scientist in reviewing (higher investment = more precise evaluation of the quality of manuscripts), as follows:

$$\hat{Q}_i^s = \alpha_j Q_i^s,$$

with α_j being drawn from a normal distribution $N(\mu = 1, \sigma = \min(T^*, Q_j^r))$, where j is the reviewer and T^* is a quality threshold which estimates the minimum amount of resources needed by each j to provide a fair review.

Suppose that the quality of manuscripts can be unequivocally quantified so that manuscripts can be compared and ranked by journals for publication. Suppose we do not consider the role of editors, the presence of multiple journals, the possibility of resubmitting rejected manuscripts and other ‘realistic’ conditions. Let us

Table 9.1 Pseudo-code of the model (for more detail, see Bianchi et al., 2018)

Input: time t , number of iterations m , set of n agents, $R, e, \tau, \Delta e, p$	
Output: publication bias, average publication quality, top quality	
1	initialize $t = 0$
2	while $t < m$ do
3	for all agents i :
4	update R_i
5	$e_i \leftarrow e_i + \Delta e$
6	compute Q_i^s
7	end for
8	while # of reviewers $< n/2$ do :
9	select random agent i
10	assign i “reviewer” role
11	match i to random j with no “reviewer” role
12	compute Q_i^r
13	compute \hat{Q}_i^s
14	end while
15	rank agents by \hat{Q}_i^s
16	for all top Pn agents in \hat{Q}_i^s ranking do :
17	published? $\leftarrow true$
18	end for
19	end while

consider these factors as irrelevant here (see the pseudo-algorithm describing the model in Table 9.1).

Let us next run our simulations for a sufficient number of iterations (in our cases, $m = 1500$) to reach a stable outcome equilibrium (in our case, we repeated our simulations at least 100 times for each initialization) and measure the outcomes as follows: (1) publication bias (i.e., the proportion of incorrectly rejected submissions on the total amount of published articles); (2) the average quality of publications; (3) average quality of the ten top-quality articles. All measurements are in time steps and so can be averaged at the end of each simulation (see the model parameter in Table 9.2).

9.3.1.1 Example 1

Let us now suppose that we want to explore a set of potential interventions to stimulate scientists to increase their quality of publication ((2)) while at the same time, minimizing publication bias at the system level ((1)). For instance, the policymaker could set up rewards or prizes to this purpose but would like to estimate the

Table 9.2 Example 1: Model parameters

Parameter	Description	Value
n	Number of scientists	500
e	Resources allocated to manuscript production	Range: 0–100% (uniform distribution)
R	Distribution of initial resources	Uniform
T^e	Minimum-quality threshold (the expected amount of resources required by each scientist to perform a review)	6
Δe	Variation of resources allocated to manuscript production	5%
P	Proportion of manuscripts accepted by journals for publication in each time step	25%

Adapted from Bianchi et al. (2018)

mediating effect of scientists’ possible reactions. You could create two ‘treatment scenarios’: one in which rewards point to strong competition and excellence, e.g., scientists are induced to compare their Q_i^s (regardless of whether their submission was published or rejected) in the top ten publications (we called it “high competition”), another one in which rewards point to the average quality (we called it “minimum expected quality”), e.g., scientists use the average quality of below-median published articles as a comparison. In both scenarios, suppose that these comparisons would determine an individual binary satisfaction value, which would make scientists revise their resource allocation decisions between investing more either in their own manuscripts or for reviewing other manuscripts.

Now, let us hypothesize three possible decisions made by scientists: (1) always selfishly investing in their own publication against peer reviewing, (2) investing more in reviewing when their manuscripts have been previously rejected, and (3) investing more in reviewing when their manuscripts have been previously published. Let us then add a control factor: a level of subjective overconfidence when scientists compare the quality of their own manuscripts with current publications by others. This can be done by re-running all the same simulation scenarios while differing for two further conditions: all scenarios initialized with ‘objective’ comparison vs. all scenarios with ‘subjective’ quality comparisons. This factorial design would imply measuring the same outcomes. Then, let us suppose that you create an artificial ‘control group’ where you remove any comparison where scientists would follow their allocation strategies without any intervention regarding ‘excellent’ or ‘minimum expected quality’ signals.

We calculated cumulative moving average values of our outcomes on the last 100 steps of each iteration and the mean value of outcome measurements for each scenario. Table 9.3 shows the first outcome ((1)), i.e., publication bias, when scientists were induced to compete for excellent or looked at minimum expected quality adapt their allocation strategies accordingly. Confront the outcomes with the control group. Adding rewards for excellence determined high publication bias than ‘minimum expected quality’ signals. However, outcomes vary greatly depending on the scientists’ adaptive reactions. Note that reviewing only after being published, e.g., a reciprocal behavior, without considering any comparison of quality was detrimental

Table 9.3 Evaluation bias (%) in different scenarios. (Mobile mean values over 100 repetitions)

Scientist behavior	Rewards				
	Control ↓	Minimum expected quality		High competition	
	Comparison bias →	Objective	Overconfidence	Objective	Overconfidence
<i>Investing only in publication</i>	57.61				
<i>Reviewing after rejection</i>	32.71	40.56	29.47	62.79	58.01
<i>Reviewing after being published</i>	66.91	27.86	28.05	30.66	27.04

Adapted from Bianchi et al. (2018)

Table 9.4 Average published quality in different scenarios. (Mobile mean values over 100 repetitions, then normalized 0–1)

Scientist behavior	Rewards				
	Control ↓	Minimum expected quality		High competition	
	Comparison bias →	Objective	Overconfidence	Objective	Overconfidence
<i>Investing only in publication</i>	0.60				
<i>Reviewing after rejection</i>	0.98	0.71	0.85	0.44	0.49
<i>Reviewing after being published</i>	0.41	0.00	0.01	1.00	0.36

Adapted from Bianchi et al. (2018)

to the publication bias. Furthermore, counterintuitively, overconfidence had a positive effect in both scenarios, especially in the high competition scenario (29.47%), where publication bias decreased even below the outcome of the ‘control group’ scenario (32.71%). Therefore, results suggest that publication bias was higher under stronger competition but precise effects depended on various behavioral factors.

If we were to consider the second outcome of interest, however, ((2), i.e., the average quality of publications), results did not vary similarly to the first outcome, i.e., publication bias. The highest value was achieved when scientists were induced to compete for excellence and reciprocated higher investment in reviewing whenever previously published (see Table 9.4). This was confirmed when considering the quality of the top ten published articles across different scenarios (see Table 9.5). In conclusion: (a) policy interventions that increase competitive spirits of scientists towards publications could backfire if norms of peer reviewing cannot be enforced; (3) even a minimal level of overconfidence can determine positive or negative outcomes compared to more objective self-evaluation (for detail, see Bianchi et al., 2018).

Table 9.5 Average publication quality of top ten published papers across different institutional settings and behavioral strategies. (Mobile mean values over 100 repetitions, then normalized 0–1)

Scientist behavior	Rewards				
	Control ↓	Minimum expected quality		High competition	
	Comparison bias →	Objective	Overconfidence	Objective	Overconfidence
<i>Investing only in publication</i>	0.51				
<i>Reviewing after rejection</i>	0.91	0.94	1.00	0.75	0.83
<i>Reviewing after being published</i>	0.36	0.01	0.00	0.93	0.34

Adapted from Bianchi et al. (2018)

9.3.1.2 Example 2

Now let us suppose that we would like to manipulate the peer-review policy adopted by journals testing the effect of shifting from confidential to open peer review in situations in which scientists would be sensitive to competition and status when reviewing others’ manuscripts. Under confidential peer review, authors and reviewers do not know each other’s identity and so they could just react to their own rejections by reducing their effort e_i in reviewing to punish the system which did not favor them. Under open-peer review, author and reviewer identities are disclosed and so scientists could reciprocate positive or negative editorial decisions by adapting \hat{Q}_i^s once they are later matched by the journal. Note that the sensitivity of scientists to this shift of the peer review model has been found in some recent ‘quasi-experimental’ analysis (e.g., Bravo et al., 2019). Do the positive benefits of open peer review come at the price of increasing publication bias, if scientists can react to status and competition and use peer review to either help favorable or punish unfavorable authors who previously reviewed their own manuscripts? Can we ideally quantify how much that price would be?

Table 9.6 shows the initial parameters of this model. We tested various possible behaviors with a focus on reviewing (e.g., always being fair, being randomly reliable, deciding how much to invest in reviewing depending on previous rejection or acceptance of their manuscript). Here, we concentrated on comparing different reviewers’ reactions to previous experience as authors in two journal settings: (1) journals following confidential peer review, in which reviewers invest in reviewing whenever previously published or otherwise disinvest, so providing unreliable reports; (2) journals following open peer review, in which reviewers and authors’ identities are revealed and reviewers reciprocate positive reviews to authors who previously favored them when reviewers, and negative reviews to previously unfavorable reviewers.

Figure 9.1 shows the first outcome of interest ((1)), i.e., publication bias, when journals follow confidential peer review and reviewers are either always fair, always unreliable, or sensitive to previous experiences as authors (e.g., being fair when

Table 9.6 Example 2: Model parameters

Parameter	Value
Number of scientists	240
Scientists' initial resources	0
Fixed resource gain (initial endowment of resources for each scientist in each time step)	1
Author bias factor (noise coefficient in the conversion of scientists' resources into quality of manuscript)	0.1
Velocity of best quality approximation (fixed rate at which the quality of submitted manuscripts varies according to the increase of the author's resources)	0.1
Discount factor on resources for unreliable reviews (discount rate on resources when scientists perform unreliable reviews)	0.5
Proportion of accepted manuscripts at the end of each time step	25%

Adapted from Bianchi and Squazzoni (2022)

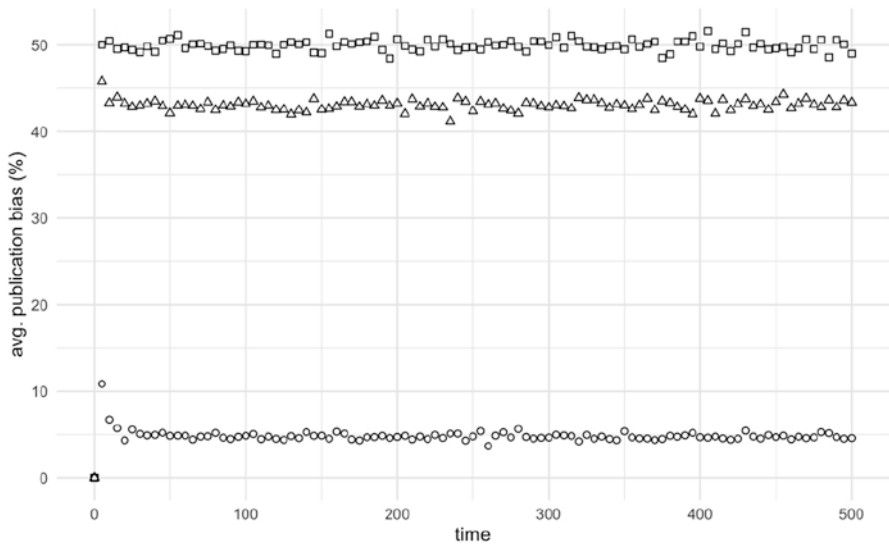


Fig. 9.1 The impact of reviewer behavior on publication bias in confidential peer review. Circles: fair; squares: unfair; triangles: reactive. Values averaged over 200 realizations. (Source: Bianchi & Squazzoni, 2022)

previously treated fairly, being unfair when previously being treated unfairly). If reviewers react to previous experience, the level of bias approximates a random situation in which the publication of manuscripts could be decided by editors tossing a coin. Let us use these outcomes as a baseline to compare the effect of reciprocity strategies in the two peer review settings.

Figure 9.2 shows the first outcome of interest ((1)), i.e., publication bias, when comparing reciprocal strategies in the two peer review settings. Publication bias increased more than 20% under open peer review and added an extra 20% of bias compared to a situation where editorial decisions would be random. This would

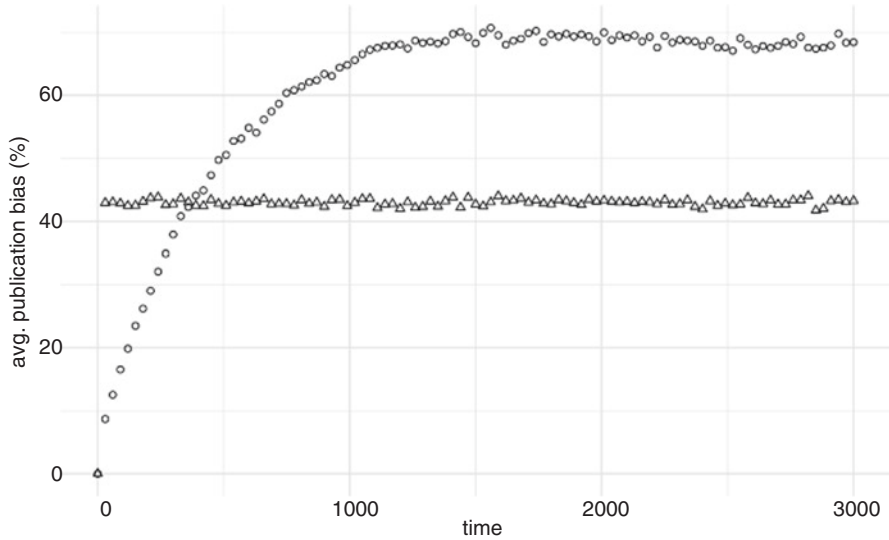


Fig. 9.2 The impact of scientists’ reciprocity strategy on publication bias in confidential vs. open peer review. Triangles: indirect reciprocity (confidential peer review); circles: direct reciprocity (open peer review). Values averaged over 200 realizations. (Source: Bianchi & Squazzoni, 2022)

suggest that open peer review could be detrimental whenever we assume that reviewers are sensitive to cooperation signals. Further results (reported in Bianchi & Squazzoni, 2022) indicate that even if reviewers would retaliate only against previous reviewers of lower academic status (i.e., with lower resources compared to theirs) while being fair in case previous unfavorable reviewers were scientists of higher status, the effect on the outcome would differ only minimally (differences not higher than 5% on the level of publication bias).

Figure 9.3 shows the effect of reviewer behavior on the second outcome ((2)), i.e., the average quality of publications. Open peer review would determine the lowest quality of publications even when compared to random editorial decisions. Note that we tested the sensitivity of these outcomes to the variation of all initial parameters and findings were confirmed (see the Supplementary Material of Bianchi & Squazzoni, 2022). In conclusion, this exercise would suggest that if practices and norms exist that make scientists frame peer review as a signaling game, open peer review policies, once adopted globally, could increase publication bias by more than 20% compared to confidential peer review, thus compromising publication quality. Obviously, other computational tests could also be designed with the model by considering for example other factors, being more nuanced, and considering empirically grounded behavior. Although a more realistic and empirically calibrated parameterization of the model would be important, as suggested by Feliciani et al. (2019) in their overview of computer simulation research on peer review, these cases here were only aimed to exemplify a method to test policy interventions artificially.

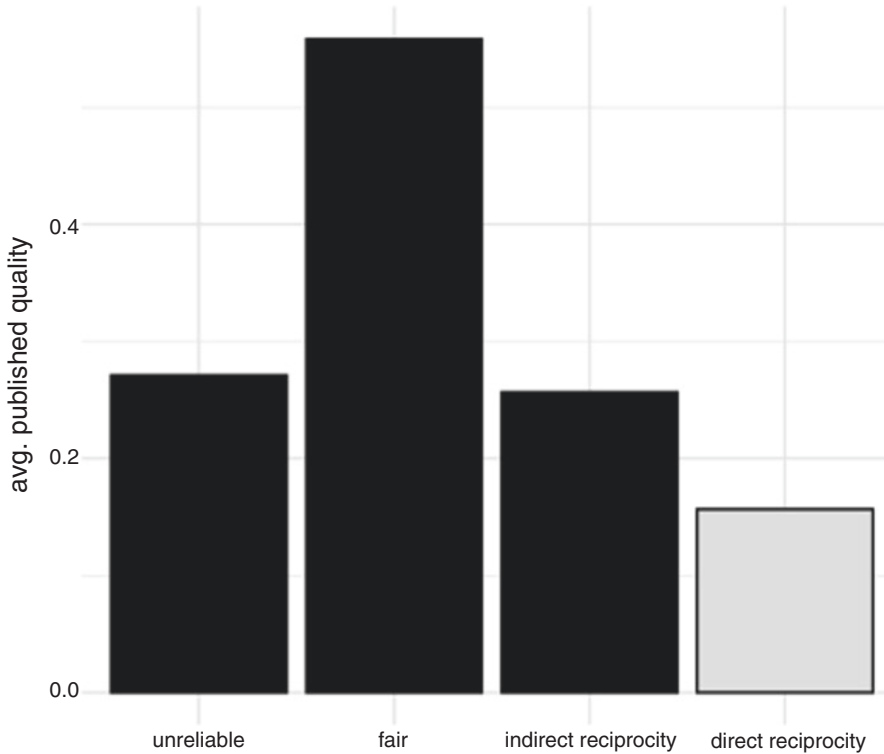


Fig. 9.3 The impact of reviewer behavior on the average quality of published papers under different peer review models. In the rectangle: comparison between reciprocity strategy in confidential (black) vs. open peer review (white). Values averaged over 200 realizations. (Source: Bianchi & Squazzoni, 2022)

9.4 Conclusions

In this chapter, we have presented ABM as a method to perform computational experimental tests on non-linear, complex effects of policy interventions as these can determine interaction effects and individual adaptations. This could enlarge the toolbox of experimental policy analysts, especially when RCTs cannot be designed due to various ethical, political, or economic constraints. *In silico* tests are also required before policy design to explore potential unintended consequences or when an understanding of social processes could provide relevant insights to enhance comprehensive policy appraisal. In our view, ABM can fruitfully complement, enrich, and even substitute—when necessary—more conventional behavioral methods for public policy.

However, the use of ABM also has important limitations. As discussed by Gilbert et al. (2018) in a comprehensive review of practices of computational modeling of public policy, deciding the appropriate model resolution requires critical decisions.

Besides the hypothetical exercises presented here, where we have proposed abstract examples, in concrete contexts, the optimal level of abstraction of a model depends on the purpose of modeling and the nature of the system being modeled (Edmonds et al., 2019). For instance, during the COVID-19 pandemic, epidemiologists have used ABMs to simulate a variety of anti-contagion policies to flatten the curve by reaching an appropriate level of resolution on certain parameters (e.g., population size). However, they followed empirically implausible assumptions on relevant others (e.g., social networks and externalities), which compromised a more comprehensive exploration of possible policy interventions while downplaying the fundamental role of uncertainty (see Squazzoni et al., 2020 for a critical overview; for an example of empirical calibration of networks in epidemiological models, see Manzo & van der Rijt, 2020).

This raises two interrelated challenges in the use of ABM for public policy, i.e., the use of empirical data to calibrate model parameters via existing or *ad hoc* data, and the heuristic value of model findings to inform policy interventions or policy evaluation (Tracy et al., 2018). In this regard, as suggested by Murray, Marshall & Buchanan (2021, 1655) in their proposed ‘target trial framework’, whenever combined with the usual experimental framework of behavioral policy, ABM could incorporate empirical data on the targeted population (e.g., calibrating salient characteristics of individuals from available data sources) and a detailed and explicit specification of the hypothetical trial, while using the *in silico* experimental nature of these models as an ‘artificial world’ “with no ethical, logistical, or financial constraints, and in which the exposure of interest is perfectly manipulable by study investigators, regardless of whether this is actually feasible or ethical in the real world.” This would help to fill the gap between empirical data and unobservable variables and inform study design. Furthermore, following Bravo et al. (2012), calibrating ABM with results from small-scale pilots, RCTs or well-detailed observational studies or re-running existing trials in a model, while scaling the characteristics of the original target population to populations with other characteristics or testing other network structures compared to those originally reproduced in the previous study, could help us to increase generalization or perform counterfactual tests of policy findings. This would help to assess the dependence of outcomes from contextual details and help us understand how much causal inference exercises on complex social behavior require careful examination.

Suggested Readings

- Epstein, J. M. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press.
- Manzo, G. (2022). *Agent-Based Models and Causal Inference*. Hoboken, NJ: Wiley & Sons.
- Page, S. E. (2018). *The Model Thinker*. New York, NY: Basic Books.

Review Questions

1. What are the limitations of RCTs for public policy?
2. What is agent-based modeling?
3. Which are the benefits of using ABM to examine social processes?

4. Can ABM be informed by empirical data?
5. What are the limitations of ABM as a method to inform policy interventions?

Replication Material

The models have been built in [NetLogo](#). The code is available at the following links:

<https://www.comses.net/codebases/6b77a08b-7e60-4f47-9ebb-6a8a2e87f486/releases/1.0.0/> (Example 1).

<https://www.comses.net/codebases/3d99eb9f-ae4f-42d0-8c58-9d28757161c0/releases/1.0.0/> (Example 2).

References

- Axelrod, R. (1997). *The complexity of cooperation. Agent-based model of competition and collaboration*. Princeton University Press.
- Battistini, E., & Bertoni, M. (this volume). Counterfactuals with experimental and quasi-experimental variation. In A. Damonte & F. Negri (Eds.), *Causality in policy studies – A pluralist toolbox*. Springer.
- Bianchi, F., Grimaldo, F., Bravo, G., & Squazzoni, F. (2018). The peer review game: An agent-based model of scientists facing resource constraints and institutional pressures. *Scientometrics*, *116*(3), 1401–1420.
- Bianchi, F., & Squazzoni, F. (2015). Agent-based models in sociology. *Wiley Interdisciplinary Reviews: Computational Statistics*, *7*(4), 284–306.
- Bianchi, F., & Squazzoni, F. (2016). Is three better than one? Simulating the effect of reviewer selection and behavior on the quality and efficiency of peer review. In L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, & M. D. Rossetti (Eds.), *Proceedings of the 2015 winter simulation conference* (pp. 4081–4089). IEEE Press.
- Bianchi, F., & Squazzoni, F. (2020). Modelling and social science. Problems and promises. In E. A. Moallemi & F. J. de Haan (Eds.), *Modelling transitions. Virtues, vices, visions of the future* (pp. 60–74). Routledge.
- Bianchi, F., & Squazzoni, F. (2022). Can transparency undermine peer review? A simulation model of scientist behavior under open peer review. *Science and Public Policy*, scac027.
- Boero, R., & Squazzoni, F. (2005). Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science. *Journal of Artificial Societies and Social Simulation*, *8*(4), 6.
- Bravo, G., Grimaldo, F., López-Iñesta, E., Mehmani, B., & Squazzoni, F. (2019). The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature Communications*, *10*, 322.
- Bravo, G., Squazzoni, F., & Boero, R. (2012). Trust and partner selection in social networks: An experimentally-grounded model. *Social Networks*, *34*(4), 481–492.
- Chalabi, Z., & Lorenc, T. (2013). Using agent-based models to inform evaluation of complex interventions: Examples from the built environment. *Preventive Medicine*, *57*(5), 434–435.
- Coleman, J. S. (1990). *Foundations of social theory*. Belknap.
- Dolan, P., & Galizzi, M. M. (2014). Getting policy-makers to listen to field experiments. *Oxford Review of Economic Policy*, *30*(4), 725–752.
- Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, *47*(4), 1–16.
- Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different modelling purposes. *Journal of Artificial Societies and Social Simulation*, *22*(3), 6.

- Edwards, M. A., & Siddhartha, R. (2017). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*, *34*(1), 51–61.
- Epstein, J. M. (2006). *Generative social science: Studies in agent-based computational modeling*. Princeton University Press.
- Feliciani, T., Luo, J., Ma, L., Lucas, P., Squazzoni, F., Marušić, A., & Shankar, K. (2019). A scoping review of simulation models of peer review. *Scientometrics*, *121*(1), 555–594.
- Forsberg, E., Geschwind, L., Levander, S., & Wermke, W. (Eds.). (2022). *Peer review in an era of evaluation: Understanding the practice of gatekeeping in academia*. Edward Elgar.
- Galizzi, M., & Wiesen, D. (2018). Behavioral experiments in health economics. In J. H. Hamilton, A. Dixit, S. Edwards, & K. Judd (Eds.), *Oxford research encyclopedia of economics and finance*. Oxford University Press.
- Giabbanelli, P. J., Tison, B., & Keith, J. (2021). The application of modeling and simulation to public health: Assessing the quality of agent-based models for obesity. *Simulation Modelling Practice and Theory*, *108*, 102268.
- Gilbert, N. (2008). *Agent-based models*. Sage.
- Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., Narasimhan, P., & Wilkinson, H. (2018). Computational modelling of public policy: Reflections on practice. *Journal of Artificial Societies and Social Simulation*, *21*(1), 14.
- Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist* (2nd ed.). McGraw-Hill.
- Grimes, D. R., Bauch, C. T., & Ioannidis, J. P. A. (2018). Modelling science trustworthiness under publish or perish pressure. *Royal Society Open Science*, *5*(1), 171511.
- Hedström, P., & Bearman, P. (2009). What is analytical sociology all about? An introductory essay. In P. Hedström & P. Bearman (Eds.), *The Oxford handbook of analytical sociology* (pp. 3–24). Oxford University Press.
- Hedström, P., & Manzo, G. (2015). Recent trends in agent-based computational research: A brief introduction. *Sociological Methods & Research*, *44*(2), 179–185.
- Hedström, P., & Udehn, L. (2009). Analytical sociology and theories of the middle range. In P. Hedström & P. Bearman (Eds.), *The Oxford handbook of analytical sociology*. Oxford University Press.
- Imai, K., King, G., & Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society A*, *171*(2), 481–502.
- Klein, M., Frey, U. J., & Reeg, M. (2019). Models within models – Agent-based modelling and simulation in energy systems analysis. *Journal of Artificial Societies and Social Simulation*, *22*(4), 6.
- Kremmydas, E., Athanasiadis, I. N., & Rozakis, S. (2018). A review of agent-based modelling for agricultural policy evaluation. *Agricultural Systems*, *164*, 95–106.
- Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, *28*, 143–166.
- Macy, M. W., & Flache, A. (2009). Social dynamics from the bottom up: Agent-based models of social interaction. In P. Hedström & P. Bearman (Eds.), *The Oxford handbook of analytical sociology* (pp. 245–268). Oxford University Press.
- Manzo, G. (2022). *Agent-based models and causal inference*. Wiley.
- Manzo, G., & van der Rijt, A. (2020). Halting SARS-CoV-2 by targeting high-contacts individuals. *Journal of Artificial Societies and Social Simulation*, *23*(4), 10.
- de Marchi, S., & Page, S. E. (2014). Agent-based models. *Annual Review of Political Science*, *17*, 1–20.
- Miller, J. H., & Page, S. (2009). *Complex adaptive systems: An introduction to computational models of social life*. Princeton University Press.
- Murray, E. J., Marshall, B. D. L., & Buchanan, A. L. (2021). Emulating target trials to improve causal inference from agent-based models. *American Journal of Epidemiology*, *190*(8), 1652–1658.

- Powell, K. E., Kibbe, D. L., Ferencik, R., Soderquist, C., Phillips, M. A., Vall, E. A., & Minyard, K. J. (2017). System thinking and simulation modelling to inform childhood obesity policy and practice. *Public Health Reports*, 132, 33–38.
- Shafir, E. (Ed.). (2012). *The behavioral foundations of public policy*. Princeton University Press.
- Squazzoni, F. (2012). *Agent-based computational sociology*. Wiley.
- Squazzoni, F. (2014). A social-science inspired complexity policy: Beyond the mantra of incentivization. *Complexity*, 19(6), 5–13.
- Squazzoni, F. (2017). Towards a complexity-friendly policy: Breaking the vicious circle of equilibrium thinking in economics and public policy. In J. Johnson, A. Nowak, P. Ormerod, B. Rosewell, & Y.-C. Zhang (Eds.), *Non-equilibrium social science and policy* (pp. 135–148). Springer.
- Squazzoni, F., & Gandelli, C. (2012). Saint Matthew strikes again: An agent-based model of peer review and the scientific community structure. *Journal of Informetrics*, 6(2), 265–275.
- Squazzoni, F., & Gandelli, C. (2013). Opening the black-box of peer review: An agent-based model of scientist behaviour. *Journal of Artificial Societies and Social Simulation*, 16(2), 3.
- Squazzoni, F., Pohill, G. J., Edmonds, B., Ahrweiler, P., Antosz, P., Scholz, G., Chappin, E., Borit, M., Verhagen, H., Giardini, F., & Gilbert, N. (2020). Computational models that matter during a global pandemic outbreak: A call to action. *Journal of Artificial Societies and Social Simulation*, 23(2), 10.
- Straßheim, H., & Beck, S. (Eds.). (2019). *Handbook of Behavioural change and public policy*. Edward Elgar.
- Tesfatsion, L., & Judd, K. L. (Eds.). (2006). *Handbook of computational economics. Volume 2: Agent-based computational economics*. North-Holland.
- Tracy, M., Cerdá, M., & Keyes, K. M. (2018). Agent-based modelling in public health: Current applications and future directions. *Annual Review of Public Health*, 39, 77–94.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

