# 2 Policy Prioritisation, Complexity, and Agent Computing

Undoubtedly, the 2030 Agenda of the United Nations' Sustainable Development Goals (SDGs) is on the right path when positing that the performance of countries involves different economic dimensions as well as issues related to social inclusion and environmental protection. However, this multidimensional perspective on development gives rise to several analytical challenges that should be addressed since, in real life, governments find it indispensable to establish a criterion for prioritising policies. The multiplicity of development traits requires governments not only to monitor a large set of indicators and establish their corresponding quantitative goals but also to consider the nuanced interdependencies among these indicators. Typically, government development plans and budgetary decisions see indicators as independent silos (i.e., without communicating vessels), but empirical evidence and theoretical arguments suggest that this view is misleading. Examples of these interdependencies are abundant, like when a minister for environment protection is concerned with the pollution that economic growth usually entails, or when a minister for education is troubled by how certain health-related issues impinge upon school absenteeism and academic performance among young individuals.

Setting punctual goals is an essential task for governments in all kinds of societies. In democratic countries, development goals may reflect social demands. In less democratic ones, they could indicate the authority's objectives or aspirations. Regardless of how a society is organised, development goals seem ubiquitous, and they should not be confused with policy priorities (a common mistake in the development and sustainability literature). Goals are a rhetorical feature of the official discourse, while policy priorities are often



https://doi.org/10.1017/9781009022910.005 Published online by Cambridge University Press

manifested through a budgetary profile. Importantly, budgets are an endogenous component of the policy-prioritisation process, while goals are (almost always) not because they emerge from longer-term processes involving several societal actors; for example, public consultations and diagnostic studies. In other words, and as stated in the Introduction, objectives without the funding needed to pursue them are just wishful thoughts. It follows that, since governments' budgets reflect their priorities, elaborating a framework to study the relationship between public expenditure and the evolution of development indicators is essential for assessing sustainable development.

There is already extensive literature on the study of the SDGs. The first group of such studies analyses specific SDGs and how micro-policy interventions affect the evolution of certain indicators (e.g., González-Pier et al., 2016; Boeren, 2019; Porciello et al., 2020; Mensi and Udenigwe, 2021; Sobczak et al., 2021). The second group emphasises the study of SDG network topology and identifies synergies and trade-offs between different SDGs or indicators (e.g., Fuso Nerini et al., 2019; Lusseau and Mancini, 2019; McGowan et al., 2019; Pedercini et al., 2019; Asadikia et al., 2021). The third group reflects on the evolution of SDG interdependencies and the possibility of transforming trade-offs into synergies in the foreseeable future (e.g., Machingura and Lally, 2017; Fader et al., 2018; Kroll et al., 2019; Amos and Lydgate, 2020; Philippidis et al., 2020). Finally, a fourth group draws on expert advice to explore indicator trends and the possibility of SDGs reaching their goals by 2030 (e.g., Ionescu et al., 2020; Luken et al., 2020; Moyer and Hedden, 2020; Pradhan et al., 2021; Benedek et al., 2021).

In spite of these valuable efforts to improve our understanding of SDG dynamics, the study of how public expenditure affects the evolution of development indicators has not been pursued in a comprehensive manner. That is to say, the available literature mainly analyses data patterns in development outcomes (networks and aggregate correlations) but misses one of the key inputs of development: government expenditure. No formal attempt exists for measuring the causal impact of specific policy instruments, such as budgetary allocations, on the observed development paths. Addressing a causal account of public funding is indispensable if we want to guide policymakers on how to meet the 2030 Agenda.

From a theoretical standpoint, the framework that we propose in this book builds on ideas stemming from complexity science and systems thinking. From a modelling perspective, such theories are formalised through agent computing. This allows us to formulate an analytical toolkit that incorporates an intricate yet realistic chain of causal mechanisms to explain the expenditure–development relationship. Because some readers might not be familiar with some of this theoretical background or with our modelling approach, our task in this chapter is to present ideas, concepts, and methods that justify our framework design. First, we briefly explain several reasons why we take a complexity perspective for modelling the expenditure–development link, and why we select agent-based modelling as a suitable tool for assessing policy impacts in development. Second, we introduce the concept of social mechanisms and explain how we apply them to measure the impact of budgetary allocations when systemic effects are relevant. Third, we compare different concepts of causality and explain the advantages of an account that simulates counterfactual scenarios where policy interventions are absent.

## 2.1 MODELLING THE EXPENDITURE–DEVELOPMENT LINK

The expenditure–development relationship, operationalised through the budget-indicators linkage, is neither straightforward nor universal since different features in the causal chain make this relationship opaque and country-specific. First, spillover effects between indicators contribute to their endogenous evolution. These contributions operate beyond the boost created through the injection of public funds into the corresponding government programmes. Hence, spillovers act as confounders in the estimation of expenditure impacts. Second, information asymmetries and uncertainty generate a misalignment of incentives between the central authority and government officials. Since the central government designs the allocation profiles and the officials are in charge of executing the associated programmes, there is a distortion in the budget-indicator link that prevents us from accurately estimating the effectiveness of such programmes, as the observed outcomes may simply reflect the incompetence of those handling the funds.

Third, structural constraints hamper the adequate working of government programmes in the short and medium term. These deficiencies are likely when infrastructure is lacking, logistic problems are endemic, human capital is poor, or ill-conceived policies prevail. Thus, estimates that do not consider these constraints or bottlenecks may incur in an overly positive perspective, or in a misguided recommendation arguing for a financial (short-term) rather than a structural (long-term) intervention. Fourth, deficiencies in public governance (e.g., the rule of law, accountability, monitoring of corruption) make it more difficult to solve collective-action and principal-agent problems associated with the proper use of public resources. Thus, not acknowledging these institutional factors may portray a smoother budget-indicator link than what it actually is.

Fifth, context matters because of the following features: countries' level of development, the degree of their institutional maturity, the nature of their indicators' interdependencies, and the setting of goals. Hence, when doing policy evaluation, one should not pool international data, as traditionally done in econometric analyses. Sixth, it is difficult to model idiosyncratic elements of a country's allocative process. Usually, this scenario happens because there are not enough data for identifying high-resolution changes in policy priorities, or because policymakers behave erratically. Seventh, country-dependent political and economic constraints (e.g., social programmes and pension funds for bureaucrats, respectively) determine the proportion of the budget susceptible to being reallocated to achieve the desired goals. In other words, there are fiscal rigidities and inertial factors that restrict the ability of a government to shift policy priorities; this should be taken into account when formulating alternative allocation profiles.

An array of analytical tools is needed to incorporate all these features in ex post policy evaluations using retrospective analyses. But most importantly, it is key to account for them when producing ex ante evaluations in prospective analyses, as this type of policy advice helps prioritise a range of government programmes. The aim of the Policy Priority Inference (PPI) research programme is, precisely, to incorporate into a computational model three general components that encompass the seven features previously discussed. These components are structural dependencies, multidimensional interlinkages, and inefficiencies induced by political economy considerations (see Figure 2.1 for a diagrammatic description). The technical challenges in formulating a parsimonious model that integrates these components lead us to combine insight and tools from complexity theory, network science, and behavioural games. In contrast to the neoclassical thinking that dominates economics, we consider that it is critical to emphasise the systemic nature of macro-level development and the high degree of uncertainty that heterogeneous agents face in the endeavour of translating public funds into policy outcomes. Accordingly, we

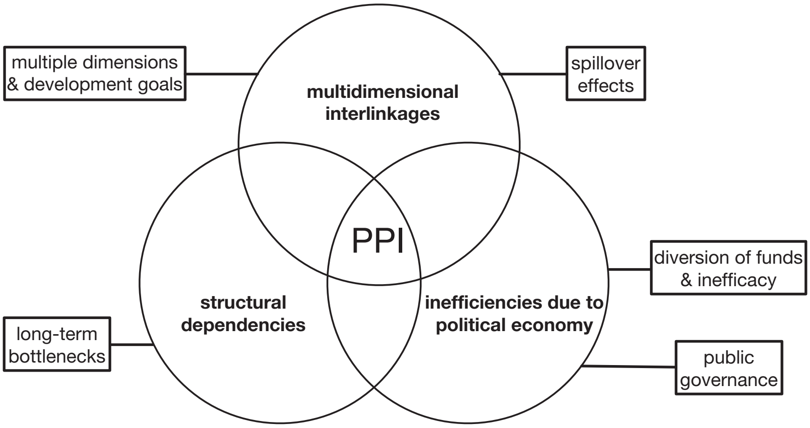


FIGURE 2.1 Factors affecting the inference of policy prioritisation.

https://doi.org/10.1017/9781009022910.005 Published online by Cambridge University Press

propose using a political economy game with learning on a network of heterogeneous agents.

We opt for agent computing since this analytical device naturally facilitates causal inference when policy interventions have the following characteristics: they occur on a recurrent basis (like periodic budget allocations), they are comprehensive (encompass multiple policy issues), they are heterogeneous across policy dimensions (allocation sizes vary substantially between policy issues), and they exert non-linear effects (their impacts might vary in terms of the policy direction, size of the intervention, or state of the country's development, for example). The account of causation underlying agent-based models (ABMs) requires discovering the data-generating process. Consequently, these models help produce prospective evaluations in which one can estimate the impact of policy innovations (e.g., alternative budget allocations to those observed in a sample period). We should add that our simulation approach builds synthetic counterfactuals from a dataset containing information from a single country. Because ABMs do not use data from other populations, they enable country-specific impact evaluations once the model has been calibrated and validated.

Economists commonly use dependency account of causation (e.g., randomised control trials, difference-in-differences, Granger causality, synthetic controls, instrumental variables, or vector autoregressions). However, there are other valid concepts of causation despite rarely being taught in the economics curriculum. In this book, we elaborate on a generative or production account of causation. Under a generative account, agent computing is used to build a procedural data-generating process (so it requires a higher detail level than more abstract statistical models). Using the specified data-generating process, one produces artificial time series with and without intervention, as explained in more detail in Section 2.2 of this chapter.

This account is not adopted only in the ABM literature; for example, some neoclassical approaches such as real-business cycles and dynamic stochastic equilibrium models also construct somehow nuanced data-generating processes.[1] But they do so by assuming equilibrium, deductive rationality, and linearity. This removes much of the procedural nature of a generative modelling framework as much of the micro-level interactions are abstracted and reduced to macro-level equilibria (i.e., a very peculiar – hardly achievable – state of the system under study). In our view, these simplifying assumptions are inadequate for dealing with the systemic considerations that arise in complex environments, in which disequilibrium is more the norm than the exception, and where non-linearity is a typical outcome of direct interdependencies between agents (i.e., aside from the reactions to changes in prices).

Another reason for the appeal of PPI, as an empirical framework, is that we can compensate for the lack of microdata with theoretical foundations and an efficient calibration procedure. In the latter, we define multiple error functions aimed at avoiding potential over-fitting problems when having a high-dimensional parameter space. Such parameters are of a structural nature, reflecting the quality of the existing government programmes and other factors that PPI does not model explicitly. Thus, these structural parameters remain unaffected by short- and medium-term decisions such as the reallocation of resources over the existing programmes. While the structural parameters are somehow aggregated, they contain information about development inertia and the returns of fiscal expenses among instrumental indicators (i.e., those associated with policy issues that receive public funds).

Despite the many elements composing our computational model, it is quite parsimonious. The model captures substantive assumptions (i.e., those with consequences on the outcomes) that are

[1] Structural causal models are another approach based on constructing a data-generating process. However, in contrast to ABMs, structural causal models do not consider multi-level relationships. See Russo et al. (2019) for a detailed explanation.

consistent with a realistic depiction of the policy prioritisation and development processes: the presence of complex interdependence structures, the learning process of public servants, the adaptation of the central authority to readjust its budget, and the political economy problems precluding a more efficient and effective use of public funds. Because our ABM is simple – but not more than necessary – it is suitable for policy guidance. However, we must be humble concerning the reach of the policy prescriptions that could be derived from its applications. Our auxiliary assumptions (i.e., those that allow closing up the model) do not allow us to tackle other important issues in the micro-level development process. For example, PPI can be used to identify policy issues with poor sensitivity to further financial assistance (potentially being idiosyncratic bottlenecks), but it is unable to explain the micro-level reasons behind such a limited response. In other words, PPI must not be seen as a universally superior tool, but rather as a complement to other frameworks that support evidence-based policymaking. In the case of unresponsive programmes, for example, PPI could be used for their initial identification, while a field experiment and ground interviews could be deployed to tease out the mechanisms behind their ineffectiveness.

Lastly, as it will be seen in Chapter 5, the model is flexible enough to deal with the fact that budgetary data from around the world are available at different levels of granularity. Some countries have only national-level budgets, others present information at the SDG level, and some others include information at the level of government programmes. The more granular the data, the lower the need to simulate how the allocation profile is formed. In the most extreme case, where data are available at the programme level, we can omit the module that specifies the government heuristic to adapt the allocation profile. Thus, the compensation of missing data by theory-driven data generation can be deployed by the user according to their particular needs.

## 2.2 GENERATIVE CAUSATION AND SOCIAL MECHANISMS

In this section, we explain the logic of vertical mechanisms underlying generative causation. The idea of this type of mechanism is intimately related to the principles behind causal chains in complex systems, in which micro-level behaviour allows macroscopic performance to emerge, and the other way around.[2] A vertical-logic tradition was originally developed by sociologists who wanted to build a tool – known as *social mechanisms* – for elaborating analytic narratives.[3] Social mechanisms refer to a set of articulated processes in which the agents (individuals or collectives) composing a system generate macroscopic phenomena through their properties, activities, and relationships.[4] According to this literature, properties refer to different pieces of the agents' behavioural rules (e.g., habits, routines, heuristics, beliefs, desires, identity, or social norms) and the information available for decision making. Activities refer to the role played by the agents as constituent parts of a system (e.g., entrepreneurs, workers, public servants, consumers, firms, or political parties), and the tasks commonly undertaken (e.g., investments, innovations, work effort, saving decisions, or consumption). Lastly, relationships describe the social, economic, and political context in which agents interact and the natural environment that conditions their way of living.

Social mechanisms link the phenomenon under study – *explanandum* – with its causal factors – *explanans* – by specifying a series of intermediate processes. Through these causal chains, it is

[2] As stated by Antosz et al. (2022), a generative definition of causation should not only be identified with explanations of emergent phenomena because it also implies the conditioning of micro-behaviours through social constructs.

[3] Robert Merton was one of the first scholars to employ the term social mechanisms in the 1960s. However, it was promoted as a formal methodology, and not only as a concept, by several posterior authors from the analytic sociology research programme: Raymond Boudon, Jon Elster, James Coleman, Peter Hedström, and Richard Swedberg, among others.

[4] For a deeper analysis on this topic, see Hedstrom (2005); Casini and Manzo (2016).

possible to connect agents with their social constructs (structures or macroscopic behaviours). These mechanisms are usually intangible since they involve processes, decisions, and social norms, but their immediate, mediate, and future consequences are observable. This sort of connection happens, for instance, when social norms condition how people dress in a community; when social pressures exert heavy influences on people's opinions; or when herd behaviour (positive feedback) produces boosts and booms in financial markets.

It is true that, in the end, individuals are the entities making decisions and taking actions; nevertheless, the context (i.e., the systems' structure) conditions their behaviour. For instance, for a new product (e.g., a technological gadget or book) to acquire wide acceptance, it requires being valued by buyers and having its appeal disseminated through social networks. Furthermore, collectives (e.g., firms, organisations, political parties) operating in a system are emergent properties (i.e., macroscopic outcomes) in a subsystem. Thus, these collectives are social constructs with their own beliefs (e.g., organisational culture) and operating methods (e.g., routines and behavioural heuristics). Therefore, they are autonomous entities and, to some degree, they are independent of their constituent agents.

Under a vertical-causation perspective, explaining social phenomena requires establishing a set of causal relationships in a macro-micro-macro circular chain (Coleman, 1986; Hedström and Swedberg, 1998), informally known as Coleman's boat. First, the deck of the boat describes how the structure or social construct evolves. Thus, the data can identify macro relationships to be explained (e.g., financial development and economic growth). Second, the bow corresponds to downward causation (situational mechanisms) that links the community structure with the agents' properties that condition the decision process. Third, the bottom of the boat's hull represents a process of lateral causation (behavioural mechanisms) in which micro entities make decisions and take actions according to their role in different markets and organisations. Fourth, the stern corresponds to an upward causation (transforming mechanisms) that

https://doi.org/10.1017/9781009022910.005 Published online by Cambridge University Press

produces macroscopic outcomes through a non-trivial aggregation of the agents' actions.

All in all, with the front, bottom, and rear of the boat, one can specify the social mechanisms that connect different periods along the systems' trajectories (boat's deck) and attempt to explain social phenomena by making them grow through agent computing. The vertical and lateral mechanisms linking micro and macro components of a complex system provide a powerful analytical machinery to study the agency-structure dilemma. Moreover, by expressing these mechanisms computationally, ABMs can integrate premises related to individual behaviour with social constructs conditioning the agents' properties. Needless to say that the mathematical methods in the toolkit of an average economist can hardly incorporate this dilemma in a comprehensive analysis and, thus, cannot explain the macro-foundation of micro-behaviour.

In Figure 2.2, we present the social mechanisms that explain the workings of PPI. For parsimony, and being consistent with our argument of goals representing aspirations rather than priorities or actions, we specify exogenous goals that the government attempts to pursue when implementing policies. We also consider structural

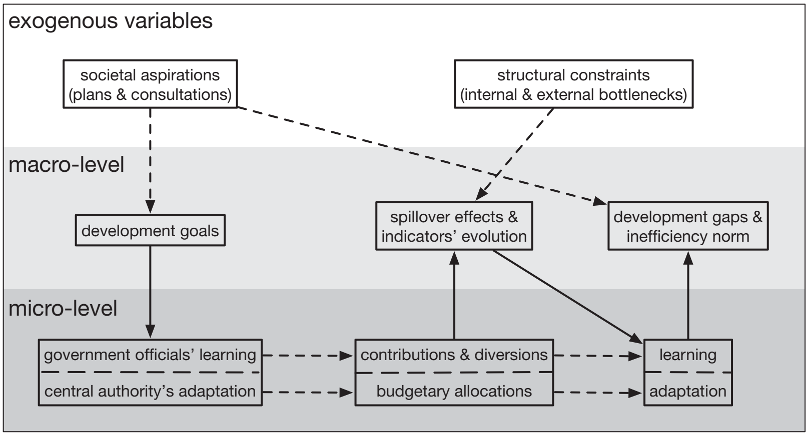


FIGURE 2.2 Social mechanisms in the expenditure–development link.

https://doi.org/10.1017/9781009022910.005 Published online by Cambridge University Press

constraints that limit the progress of development indicators in the short- and medium-terms. Although these goals are set, in part, by societal aspirations, especially in democratic regimes, in real politick, they also reflect political influences of lobbying groups and strategic manoeuvres of elected officials and technocrats who want to show their proficiency by producing results. In the case of long-term constraints, there are two kinds of bottlenecks: *internal* and *external*. Internal bottlenecks consist of ill-conceived government programmes that make indicators unresponsive to increases in public funding. External bottlenecks produce inertia that retards the indicators' evolution because of problems related to the context where these programmes are implemented, such as the lack of adequate infrastructure. In this book, we do not study policy interventions related to structural issues since we only aim to measure the impact that a change in the budgetary profile or other funding sources has on development indicators.

We start the chain of social mechanisms at the macro level, with data showing worrisome development gaps, possibly due to structural considerations and a historical lack of funding. Then, at the micro level, the central authority of a new administration decides to react by adjusting the budget allocated to a set of programmes (downward causation). At the same time, public officials implementing those programmes learn, through their previous experiences, how much of the received resources should be employed – as mandated – to accomplish their own objectives. As a consequence, there is a waste of public funds due to administrative inefficiencies or the diversion of resources for personal use. These decisions consider the socioeconomic context and the performance of public governance institutions monitoring and punishing the inappropriate use of public funds. Then, they translate into actions, and a fraction of the budget becomes the contributions that functionaries employ to implement policy issues (lateral causation).

Yearly funding creates differentiated progress in development indicators (upward causation). Such progress is also shaped by interdependencies and spillover effects at the macro level of the system. The cycle repeats several times and, in the process, a social norm of inefficiency settles at the macro level, which conditions how functionaries act (i.e., a standard of high inefficiency makes them less worried about misbehaviour because it seems socially acceptable). At the end of the administration, budgetary decisions can be successful or not in closing development gaps. Consequently, the computational model helps us to estimate policy impacts through synthetic counterfactuals. These counterfactuals produce what-if scenarios (of absent interventions) assuming the same data-generating process where policy interventions occur factually, as explained in the next section.

## 2.3 ON CAUSAL INFERENCE AND AGENT COMPUTING

One of the fundamental challenges for social scientists is to discover and estimate causal impacts in the phenomena under study because available data usually come from non-experimental settings. Understanding how exogenous shocks or policy interventions impinge upon critical variables is essential to tame the systems' evolution and to avoid undesired outcomes. From the natural sciences, we know that causal inference requires comparing an outcome variable in an intervened system with the performance of such variable if such intervention would not occur. In other words, identifying and measuring the magnitude of a causal impact requires making comparisons with counterfactuals (i.e., facts produced in the same system but without interventions).[5] Hence, if experiments (controlled settings) are unethical or unworkable (i.e., difficult to design or implement), then the main problem of causal inferences is finding a proper counterfactual.

Heckman and Pinto (2022) synthesise the connection between causation, counterfactuals, and impact evaluation along the following lines:

[5] In a popular book, Pearl and Mackenzie (2019) argue that counterfactuals are the main distinguishing factors between causal and non-causal statistical analysis.

> Good policy analysis is causal analysis. It analyses the factors that produce outcomes and the role of policies in doing so. It quantifies policy impacts. It elucidates the mechanisms producing outcomes in order to understand how they operate, how they might be improved and which, if any, alternative mechanisms might be used to generate outcomes. [. . .] It systematically explores possible counterfactual worlds. It is grounded in thought experiments – what might happen if determinants of outcomes are changed. [. . .] Models and thought experiments are central to economic analysis.

These authors point out that one cannot perform counterfactual experiments without making explicit reference to an account of the causal mechanisms underlying a system. In social sciences, there is no consensus on the appropriate way to conceive causality. We can group the existing epistemological approaches into two main categories.[6] The first consists of a dependency account in which causal factors make the existence of one or more consequences possible (i.e., effects appear). The second takes a production (generative) account in which causal factors help to generate consequences (i.e., effects grow).[7]

These two approaches employ alternative mechanisms to determine causal effects. Such mechanisms refer to the type of links analysts use to relate variables in data. The horizontal mechanisms, common to dependency accounts, establish a system of relationships among variables defined with the same level of aggregation. Such a system specifies chains or networks of variables linking interventions with their consequences. The vertical mechanisms, common to the production account, posit causal relationships at different levels of

[6] For more details, consult Hall (2004); Casini and Manzo (2016).

[7] In the dependency account, causality comes from the following criterion: $X$ causes $Y$ when manipulating the first variable affects, systematically, the second one, controlling for a set of confounding variables, $Z$, possibly related to $Y$. Under the production account, a causal relationship exists when the following criterion holds: $X$ causes $Y$ when starting with conditions set by the first variable, there is a 'generative process' (i.e., a collection of interlinked procedures) that can produce the information underpinning the second variable (i.e., it can replicate statistical regularities).

aggregation, in so far as actions at the micro-level have an incidence on macroscopic behaviour and the other way around.

Therefore, the applicability of these two schemes depends on the nature of the data and the problem at hand. The dependency account is suitable when the intervened variable and the outcome exhibit the same level of granularity (i.e., micro-level studies in which the units of analysis are individuals, organisations, or firms.) These may include, for example, data at the school level to study whether health-related absenteeism impinges upon the students' average academic performance in standardised tests; and data at the individual level to estimate the causal impact of the parents' schooling level on their offspring's social status.

The generative account is convenient when the intervention policy occurs at a low level of granularity (e.g., countries, firms, or schools) but its potential effects have consequences on a higher scale (e.g., a country). For example, when studying how financial development relates to countries' economic growth – both aggregate variables – one has to identify causal factors at a lower level of granularity than the outcome variable. That is to say, intervention policies (e.g., regulatory schemes and competition rules) directly affect units operating on a lower scale (e.g., banks and other financial intermediaries). This conception of causality does not rely on econometric tools since it requires a systemic approach where it is possible to explore spillover effects and the emergence of social phenomena.

### *2.3.1 The Identification of Counterfactuals*

In the past, when academics identified a policy intervention, they frequently used output data from the pre-intervention period as a counterfactual. In their mind, this approach was valid since the data came from the same system. However, comparing outcome variables – before and after the intervention – is not sufficient to determine the existence or significance of an impact. Most quantitative social scientists know by now that this approach is misleading since the facts reflected in the outcome variable during the pre-intervened

https://doi.org/10.1017/9781009022910.005 Published online by Cambridge University Press

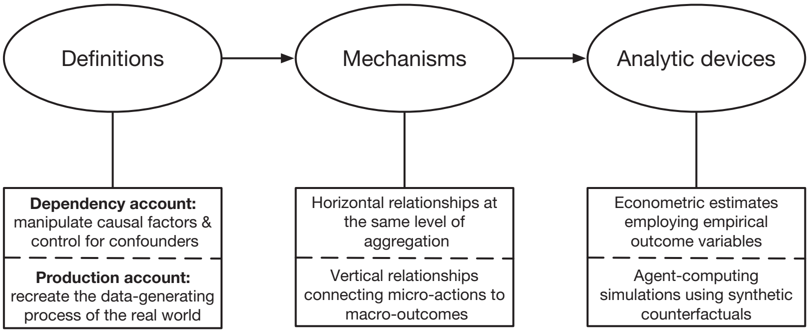


FIGURE 2.3 The epistemological triad of causation accounts.

period are not an adequate counterfactual. Formally speaking, this approach disregards ex ante, the potential set of additional variables, unrelated to the shock or policy, that affect the inter-temporal changes observed in the system during the entire sample period. For instance, a steady fall in employment due to some intrinsic property of the system might bias the impact measurement of a shock when employing only inter-temporal comparisons. Furthermore, in a non-linear setting, a downward trend in employment might appear only after the shock; hence, a comparison with a projection of the pre-intervention data would be erroneous as well.

In recent years, nuanced techniques have appeared to identify counterfactuals and, in this manner, to make valid comparisons with adjusted data to reflect 'what-if' scenarios.[8] Here, we briefly review two alternative analytical devices for proposing counterfactuals (see Figure 2.3). Each alternative makes different theoretical assumptions and looks at different facts in the real world to produce data points or time series that can function as counterfactuals.

When implementing a dependency account of causality, analysts look mainly at changes in the performance of the outcome variables, and econometric estimates are controlled for time and fixed

[8] For textbook explanations of econometric methods of causality, see Angrist and Pischke (2009); Dunning (2012); Cunningham (2021); Hernan and Robins (2023); for a formal overview of these methods and their applications, see Athey and Imbens (2017); de Chaisemartin and D'Haultfœuille (2020); Roth et al. (2022); Abadie (2021).

effects. They collect outcome data for different periods (pre- and post-intervention) and populations (with and without intervention). As mentioned previously, these studies use information with the same level of aggregation, either micro or macro data. The aim is to have information from similar populations, which allows the implementation of a 'natural experiment'. The populations from which the data are collected are supposed to correspond to systems that only differ by the presence of an intervention.

In contrast, when implementing the generative account of causation, analysts look at a wider set of information for describing the context and outcomes where a population operates, which enables the elaboration and calibration of a computational model. Under this approach, qualitative and quantitative data are collected at different levels of aggregation (micro, meso, and macro) to produce artificial data for the outcome variables with and without interventions. In other words, empirical data of different kinds help build vertical causal mechanisms, specifying exogenous variables, calibrating initial conditions and parameters, and validating the model. Therefore, the objective is to discover the data-generating process of the system under analysis and to simulate non-intervened series for the whole sample period. The simulation of non-intervened series fit, by construction, the requirements for being a counterfactual.

This succinct description states that each method has different theoretical underpinnings and data requirements. Consequently, their applicability depends on the availability of certain forms of data and the nature of the working hypotheses.

### *2.3.2 The Workings of the Dependency and Generative Accounts*

Generally speaking, the dependency framework requires finding empirical systems (populations) similar to those affected by shocks or policies. When conditions characterising both types of systems are not identical, it is possible to control for the effect of non-observable variables (e.g., individuals' intelligence or motivations) and observable ones (e.g., individuals' sex, parents' education, and place of birth) by taking differences in the output data in each of the two systems across time – before and after the intervention; a procedure that eliminates biases due to time-invariant variables affecting each system in a differentiated manner. Unfortunately, correcting for such a bias is not enough, since there could be variables whose value might change over time in both systems. Therefore, an additional adjustment in the statistical analysis of the data ensues. One way to control for these time-varying variables consists of including confounders that, in regression analyses, produce alternative dynamic paths. Consequently, to calculate the impact of an intervention (i.e., another time-varying variable affecting both systems in a differentiated manner), the recommendation is to adjust the data once more through differences in the inter-temporal variation of each system (i.e., the first differences). The second difference eliminates the time effects of unknown variables when their values are identical in both systems.

There are two main technical problems with a dependency account when applying difference-in-differences (DiD) estimates for impact evaluation. First, the assumption of 'parallel trends' between the intervened and the counterfactual systems. Second, selecting the suitable econometric equation to eliminate biases created by the omission of relevant variables. The parallel trends assumption is not falsifiable since it requires that both systems operate with the same intrinsic dynamics, either if it experiences policy interventions or not (i.e., the non-recognised time-varying variables affect both systems similarly). This task is impossible to verify since we cannot know how the system dynamics work in a scenario of no intervention (a counterfactual). At most, we can check, statistically, if the dynamics are similar, just in the pre-treatment period, for the outcome variables in the factual and counterfactual scenarios (e.g., if they present a similar trend). However, this test is not enough to fully validate the assumption.[9]

Concerning the econometric formulation to estimate the parameter of a causal impact (i.e., interaction term identifying the intervened systems during the post-treatment period), it is not always enough to control for time and fixed effects (i.e., setting differences in the output variable between systems and periods, respectively), and confounders. The equation needs proper identification to avoid additional biases. The selection of explanatory variables and functional forms implies checking different formulations: assessing if the dependent variable presents an auto-regressive process because of lagged impacts; if there is a spatial correlation between the different types of systems used in the dataset (i.e., there could be spillover effects from an intervened system to a counterfactual one); or if there are non-linear impacts that depend on the starting values of the intervened and outcome variables.

Figure 2.4 indicates that, in order to establish a counterfactual, the generative framework does not require gathering information from other populations. Instead, it is necessary to dig deeper into the causal mechanisms of the population under analysis to discover the data-generating process of the real world. Once the ABM reflects the context of the system and satisfies validation tests, we can use it to generate artificial data for the outcome variables in both scenarios: with and without policy intervention. In this methodology, the time series simulated when turning off interventions are, in principle, good candidates for counterfactuals. By construction, these data come from the same process as the one underlining the intervention scenario.[10] However, as we explain in Section 2.3.3, there are some caveats for considering that, irrespective if the intervention is active or

[9] Placebo tests are advisable; for instance, checking if random points in the pre-intervention period indicate the absence of a causal effect.

[10] The most appealing property of agent-based modelling is its ability to allow for what-if analyses (counterfactuals) when real-world experiments are not feasible (Herd and Miles, 2019).

https://doi.org/10.1017/9781009022910.005 Published online by Cambridge University Press

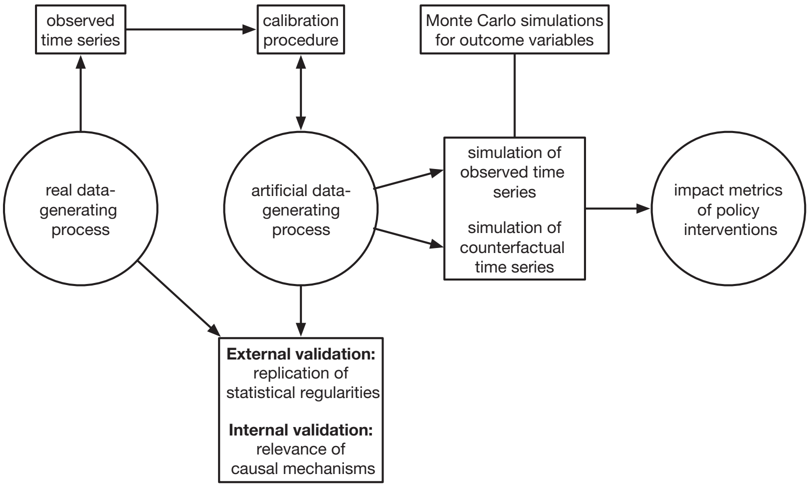


FIGURE 2.4 Causal inference through agent computing.

not, the ABM's simulated data can be empirical descriptions from a similar system. Likewise, it is worth mentioning that the actual data-generating process is not necessarily the same as the artificial data-generating process. Hence, the quality of a causal inference depends on the quality of the ABM's calibration procedures and validation tests.[11]

In Castañeda and Guerrero (2019), we argue that ABMs have certain features that help them prevent the shifting of parameters between the calibration and the evaluation stages. First, a bottom-up perspective allows the inclusion of micro-level parameters (e.g., technology and risk aversion) that are relatively more stable than those describing macro-level relationships (i.e., they are less responsive to changes in incentives and different macroscopic conditions).[12]

[11] For an extended discussion on this issue, see Manzo (2022).

[12] However, it is not always the case that a micro-parameter is invariant to changes in incentives. For instance, the literature on behavioural economics has shown that some material incentives might crowd out pro-social behaviour (Bowles and Polania-Reyes, 2012). Hence, if the preference separability condition does not hold, then the modellers would have to use a functional specification for allowing this possibility as a result of policy interventions.

Second, incorporating a micro-macro causal chain makes it possible to replace fixed macro-parameters with social constructs (e.g., social norms, preferences, behavioural heuristics), which are endogenous in these models. Third, the flexibility of the simulation approach makes sensitivity analyses possible. These procedures help verify whether reasonable variations in parameter values exert a significant influence on the results. Therefore, ABM simulations are meaningful for guiding policies since they are less prone to the Lucas critique of macro-econometric analyses (Lucas, 1976).

In the case of PPI, the parameters that require calibration refer only to long-term structural constraints. This feature increases the reliability of our model for producing counterfactual simulations since these parameters are not attached to incentives that might be affected by budgetary changes. This independence holds because these parameters describe meso-relationships between budgetary expenses and the dynamics of the outcome variables (development indicators). Moreover, the value of these parameters (and a spillover network that is a model input) come from country-specific data. An estimation procedure that relies on these data implies that the parameters describe a specific reality. Hence they do not represent average behaviour, like those coming from regression analyses using panel data. However, the reader should be aware that these parameters characterise the development process of a country's past. Accordingly, we assume that the government programmes and their historical operation should not differ substantially from those that will prevail in the short- and medium-term.

The shortcomings of adopting a generative approach to causal inference refer to time constraints and the technical difficulties of building computational models. This endeavour requires understanding the vertical mechanisms connecting the micro-intervention with the macroscopic outcome. Nonetheless, it offers several key benefits, such as the possibility of studying policies that have not occurred empirically. At first glance, a shallow critique would argue that the generative account is more demanding in terms of data; however, this is not always the case. For instance, econometric studies for producing DiD's estimates require panels with populations of the two types of systems, or cross-sections with two points in time (pre- and post-treatment), while computational simulations only need historical data of one population. Furthermore, econometric data has to be perfectly structured in time and space to estimate regression coefficients. In contrast, ABMs can work with stylised facts of initial conditions and exogenous variables, and use bespoke metrics of endogenous variables for calibration purposes. In any modelling framework, better data will facilitate producing better models. In the case of an agent-computing model, better data help to include additional validation tests at different levels of aggregation, thus improving its reliability.

### *2.3.3 The Validity of Agent-Computing Counterfactuals*

A counterfactual observation is an outcome that could have been realised (or grown) in a hypothetical setting (i.e., contrary-to-fact) of a system that, instead, experienced an actual fact (e.g., an intervention). Therefore, establishing a counterfactual implies 'rerunning history in a system, but without the presence of the actual fact'. This task is unfeasible when referring to the real world because of the fundamental problem of causal inference (Holland, 1986), in which one cannot observe the performance of a particular system (or unit of analysis) with and without intervention at a specific point in time. Hence, the only way to produce a counterfactual is by utilising a thought experiment where we can manipulate the states of a system by 'turning off' an intervention. At first sight, it seems that ABMs are helpful technical devices to deactivate interventions and 'rerun history' since they can generate artificial data under different environments of the same system.

Davis (2018) suggests that, due to the use of a data-generating process, simulation models are ideal tools to recreate outcomes from similar systems that only differ in the presence or absence of a causal factor. The social mechanisms in an ABM can resemble those prevailing in reality, and remain unaltered when the modeller tweaks the system by modifying a set of parameters or exogenous variables. This feature is, undoubtedly, a first step in the right direction for assuring the similarity of the system when rerunning history several times. However, when agent computing aims to model a complex system, some complications may arise, precluding a straightforward interpretation of the simulation results. One issue relates to the sensitivity to initial conditions (non-ergodicity) that exists in some complex systems. When a model is sensitive to initial conditions, it is not evident whether divergent paths are a consequence of an ulterior intervention or if they are being affected by different initial conditions.[13] Another issue has to do with complex systems being prone to transition phases and positive feedback that, on some occasions, might alter the trajectory of the systems when certain contingencies accumulate.[14] Again, divergent paths may ensue from a series of random factors, igniting a tipping point or a sharp change in the outcome trend. Although the existence of these non-linear effects is independent of the intervention state (active or not), their mean size and variability might not be.

The aforementioned complications may preclude ABMs from fulfilling the similarity condition. Nonetheless, we can handle the problem of non-linear random effects (initial conditions and cumulative contingencies) through Monte Carlo simulations.[15] The idea of having several repetitions of stochastic simulations is similar to field experiments using randomised controlled trials, where

[13] Marshall and Galea (2015) advocate for using agent computing, especially when there are multi-causal channels and interference (i.e., spillover effects between agents). Yet, they also warn of non-well-defined causal effects when non-ergodic patterns prevail.

[14] Building on the ideas of Arthur (1990) and Elsner (2017), DeMartino (2021) argues that, in complex systems, the world tends to be historically contingent, and that processes of relative stability might alternate with sudden episodes of volatility.

[15] Herd and Miles (2019) present another simulation method for finding valid counterfactuals when testing for 'token causation'. Under this proposal, analysts attempt to decipher if a particular event in some period is a causal root of another event occurring in an ulterior and specific period of the same process. This type of causation is not of concern here. In this book, an intervention is not an event that becomes active at just one point in time, but an exogenous sequence of events that alters the environment during all periods included in the analysis.

micro-units for the treatment and control settings are selected randomly. Randomisation aims at "cancelling out" the influence of factors that are different from the treatment. However, instead of just paying attention to mean estimations of the impact metric (e.g., differences with respect to the baseline simulation), we must look for non-causal divergent paths by revising the distributions emerging from the simulated data. When an ABM generates a one-mode distribution of the relevant impact metric, and it exhibits small tails (e.g., Gaussian, bell-shaped, or exponentially decaying), the non-linear elements of the system are unlikely to create invalid counterfactuals (i.e., simulated data describing different populations). In contrast, when the simulated data generate a distribution with multiple modes or significant extreme events, a more cautious interpretation of the results is in order.

Figure 2.5 includes two examples of distributions of an impact metric where inadequate counterfactuals are likely to emerge. Figure 2.5a presents the case of a bimodal distribution. Such an outcome could be produced by a high sensitivity to initial conditions. In this example, the first hump indicates that the causal impact is negative, while the second one shows a positive one. This contradictory result could be avoided by rerunning the model with the same random seed at the beginning of the simulation. Accordingly, the same stochastic realisation materialises all along the simulated

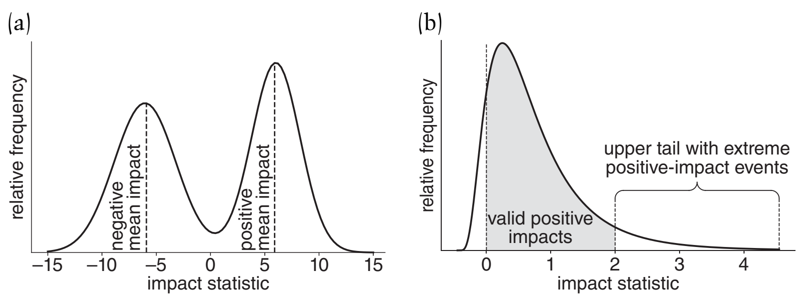


FIGURE 2.5 Causal impact distributions with inadequate counterfactuals. (a) Sensitivity to initial conditions and (b) overreaction due to non-linear random effects.

period in both scenarios (with and without intervention), even if in some of these runs the presumed causal event has been 'turned off' at some point for the rest of the simulation. Nevertheless, the potential problem might not appear in a specific example when rerunning history – with independent Monte Carlo simulations – produces a unimodal distribution of the impact metric. Importantly, the initial conditions should reflect reality for the procedure to be valid.

Figure 2.5b presents the case of an impact metric with extreme events in the upper tail, indicating a lower probability of realisation – but not negligible, like in heavy-tail distributions (e.g., Pareto, log-normal, or Weibull). These unlikely events might be the consequence of an intervention where non-linear random effects not only increase the mean value of the outcome variable (i.e., the causal effect) but also the probability of larger differences when comparing opposite realisations of the intervened and baseline simulations (i.e., an overvalued impact). A suggestion would be to identify such tails – on either side of the distribution – and trim them before calculating *p*-values for assessing statistical significance.[16] Presumably, tail events produce invalid counterfactuals because they describe different systems when working with and without interventions (i.e., they are false positives since they do not reflect either positive or negative causal impacts).[17]

Another, and more promising, alternative is to rerun the model several times in pairs of trials (i.e., with and without intervention) but controlling for the simulations' stochastic variations. In practice, this procedure consists in employing the same random seed along the simulation run to guarantee identical shocks in both simulation variants. When applying the random seed at the initialisation stage,

[16] Another option is to substitute extreme differences for other pairs of simulation runs that do not exhibit such differences.

[17] Unfortunately, trimming extreme values of these comparisons runs the risk of removing false negatives too from the distribution of impacts. They could be erroneously discarded as valid comparisons, even when the associated counterfactuals describe large impacts that do not come from cumulative random effects moving in opposite directions. However, in these tails, we would expect valid counterfactuals that are less common than invalid ones (i.e., those coming from divergent random shocks).

as suggested by DeMartino (2021), one preserves agents' properties (e.g., preferences, risk aversion, wealth, or technological know-how) and initial interdependencies. But this random-fixing scheme is not enough since we also need to assure that the same cumulative non-linear random effects, if occurring, take place in each pair of trials across the complete sequence of social mechanisms. In other words, the counterfactuals are valid when forcing the ABM to produce similar sequences of random events as the simulation proceeds. Consequently, regardless of the random patterns of macroscopic behaviour that might emerge when rerunning history, the fixing of the model's stochastic components produces worlds that look alike on pairwise comparisons.[18]

Finally, we should mention that non-stationary data patterns are not rare in agent-computing models. A consequence of non-stationarity is that statistical metrics of the outcome variables could change over time. Nevertheless, this issue does not imply a violation of the similarity condition. Instead, it is a virtue when disentangling the short- and long-term effects of interventions.[19] For instance, an impact metric with different distributions over time allows for the size of the causal effect to exhibit variations in the temporal dimension.

### *2.3.4 The Benefits of Using Agent Computing for Policy Evaluations*

Causal inference with computational models is not always a better alternative than measuring causal impacts with econometric tools. This statement holds when using microdata for the intervention and

[18] Non-linear effects in the outcome variable (e.g., discontinuities) are not necessarily the result of random factors creeping in but also a direct consequence of the intervention itself. Since the deterministic components of the model's social mechanisms capture these effects, the associated causal impact will frequently appear when doing Monte Carlo simulations (i.e., they will not form part of the tails in the distribution of these impacts).

[19] However, non-stationary data could be troublesome for some calibration procedures. For example, the length of an empirical time series may give place to different metrics and, hence, alternative parameter estimates. For further details, see Grazzini et al. (2013); Grazzini and Richiardi (2015).

outcomes variables, or when the macroscopic system is cumbersome. The latter feature makes extremely difficult the modelling of the real data-generating process, either because of information limitations or due to a lack of knowledge precluding the identification of empirically relevant mechanisms. This scenario might happen when interventions have widespread ramifications across the economy and society, or when they exhibit a mixture of short- and long-term impacts. If this is the case, a dependency account may be the only option available since it does not require deciphering the chain of mechanisms linking the policy with the outcome variables.

On other occasions, we advise producing the two types of policy evaluations and looking for coincidences when the aim of such studies is policy formulation; for example, as we do when studying the impact of international aid in Guerrero et al. (2023), or when we compare different expenditure–impact assessment frameworks in Guariso et al. (2023b). Nevertheless, one should consider that, with the ongoing information-technologies revolution, and as big data becomes commonplace, it is becoming more common to identify social mechanisms across the multiple scales of socioeconomic systems, so agent computing is gaining a reputation as a reliable analytic tool. Finally, we would like to discuss ten reasons why agent computing can be more suitable for producing causal inferences. The reader should be aware of these advantages when weighing the pros and cons of different methodologies.

1. **Empirical evidence is not available**: Analysts or policymakers can apply agent-computing models to study the impact of innovative policies in specific contexts. Sometimes, theory suggests policies that have not been implemented in the past, either locally or somewhere else. However, before implementing them in reality, evidence-based policy recommendations require producing ex ante evaluations of their potential impacts. This endeavour is feasible with ABMs since intervention scenarios and counterfactuals come from artificial data.
2. **The context is different**: Although on many occasions, it is possible to find empirical evidence on a specific policy intervention in the academic literature, it might not be adequate to extrapolate those results to a

different context, either because the size of the intervention varies or because the environment is quite distinct. Therefore, we can use a computational model to analyse the different sizes of interventions, to re-frame them, or to consider alternative contexts. Accordingly, sensitivity analyses with ABMs are equivalent to meta-studies of econometric estimates using different datasets.

3. **A natural experiment is not possible**: Agent-computing simulations can be a good option when one cannot identify a natural experiment. This situation might hold if we cannot find enough non-intervened systems that are likely to operate as proper counterfactuals. There are different reasons why this scenario may occur: (1) when we reject the parallel trends assumption in most available candidate populations; (2) when few populations can operate as a counterfactual system and, thus, one cannot produce reliable statistical estimations with econometric tools; (3) when policies generate spillover effects in all tentative populations; (4) when there are heterogeneous policies in the populations included in the set of the intervened systems; (5) when a policy operates at an aggregate level and, thus, no untreated systems are available for comparisons.
4. **Building synthetic controls is not advisable**: When there are no suitable candidates for counterfactuals, econometricians suggest using synthetic controls. With this methodology, one can construct an artificial outcome variable (e.g., growth or GDP per capita) with a dataset composed of 'donor' populations (e.g., countries). A synthetic control is acceptable when it replicates the trajectory of the real variable along the pre-intervention period. Then, it is possible to produce an impact assessment by comparing the real output data with the projection of the synthetic control. However, this estimation is unreliable if some of the populations in the set of donors exhibit spillover effects from the intervened policies; or if some of these donors also experience their own interventions during the post-treatment period and the covariates that produce the synthetic control cannot capture such effects.
5. **Multiple interventions within a short period**: In real life, it is common to find that several policy interventions happen simultaneously or within a short period of time. In this setting, it becomes more difficult to find adequate counterfactual populations that, at least, do not present one of such interventions. In contrast, in an ABM, one can generate artificial data assuming one or several of these interventions when a model of the data-generating process is available. Hence, with this

methodology, we can estimate the accumulated impact of each intervention in the system (e.g., having a spur of public expenditure through several short-lived government programmes).

6. **Policy refinements**: A proposal of policy changes is frequently the result of an existing policy incapable of generating the expected outcomes. Under these circumstances, a policy reformulation is in order, and for this, we need to understand how causal mechanisms operate. Therefore, we can use agent computing as a virtual laboratory to discover which policy refinement might affect, in a positive manner, the outcome variables under analysis.
7. **Measuring impacts for specific populations**: DiDs and other econometric tools require a set of treated populations; thus, the estimated impact corresponds to the hypothetical average of the systems composing the dataset. However, when the aim of a study is to support policymaking, it is convenient to assess the impact of a specific system. One can use a synthetic control approach since no other system exhibiting the policy intervention is required, but another possibility is to formulate the assessment through ABM simulations. The latter has the added advantage that simulated counterfactuals describe the context where the system operates and preclude the existence of spillover effects from the policy to be analysed.
8. **Allowing non-linear and dynamic impacts**: Traditionally, econometric analyses assume that impacts are linear. Thus, they are invariant to the magnitude of the intervention, the direction of the changes, and the size of the outcome variables. When dealing with complex systems, the opposite scenario tends to occur. Hence, we might consider analytical tools, such as agent computing, that allow for this possibility. In contrast, to measure dynamic impacts with econometric tools, several parameters have to be included when formulating the regression equation. However, this option is not viable with a relatively short time series (a common feature in development indicators). Instead, we can formulate an ABM that, by construction, is dynamic and, hence, produces synthetic time series capturing the complete evolution of the observed impacts.
9. **Interventions are not truly exogenous**: In most cases, governments implement policies to prevent undesired outcomes that are likely to happen or are already taking place (e.g., the government might increase a value added tax to avoid fiscal unbalances prone to inflationary spurs).

The conditions for these outcomes might not be present in those systems selected as counterfactuals in DiD estimations, even if the parallel trend assumption is not rejected in a pre-intervention analysis. In contrast, we can employ the synthetic counterfactuals of computational models since, when properly calibrated, they already contain the seeds of these undesired outcomes. In other words, both types of artificial data, intervened and counterfactual, are comparable since they share all factors underlining their dynamic processes but the policy (see Geanakoplos et al. (2012) for an example with a model of the US housing market bubble of 2008 and counterfactual simulations of interventions by the Federal Reserve).

10. **Suitable impact metrics**: Most econometric tools only measure average impacts at a point in time or in a sequence of periods for a specific outcome variable. On the contrary, in agent computing, analysts can design bespoke impact metrics that capture the whole non-linear dynamics in the trajectories of the simulated series, intervened or otherwise (as we do in Chapter 9). By measuring the gap between these trajectories across time, we can obtain a more detailed estimate of the accumulated impact up to any cut-off point. Furthermore, with many outcome variables classified by activity, sector, or region (i.e., GDP), a pre-designed metric allows measuring impacts with different levels of aggregation, from one variable in one activity and region to all variables across all activities and regions.

## 2.4 SUMMARY AND CONCLUSIONS

In summary, the problem of policy prioritisation and, more generally, of sustainable development conveys significant analytical challenges. PPI builds on different theoretical traditions about how social systems operate and harnesses the power of agent computing to operationalise such ideas. With this framework, it is possible to overcome some of the limitations of more traditional causal–inference methods and to provide reliable tools to support evidence-based policymaking. In trying to provide arguments as to why ABMs are a reliable alternative analytical tool, we discuss ten reasons a researcher should consider when using agent computing for solving a causal-inference problem.

However, taking advantage of this and other similar frameworks requires an open mind about how we conceptualise the problem under study and a certain degree of computational literacy, given that the latter is not common among social scientists. Thus, this chapter should also serve as a call for action for a much-needed updating of social science curricula, both in terms of quantitative methods and interdisciplinary theoretical frameworks. We elaborate on this need in Chapter 13.