4 A Computational Model

Now that we have motivated the use of agent computing and elaborated on methodological and empirical challenges, it is time to introduce the model that underpins much of the analysis conducted in the PPI research programme. As PPI has developed, our model has evolved, gradually improving theoretical and empirical features that are important to tackle development-related problems. Thus, the model that we employ in this book is the latest version, which we developed in Guerrero et al. (2023).

Our model describes a government that allocates budgetary resources across several policy issues (development dimensions) and functionaries that, through government programmes, transform public spending into policy outcomes.[1] Because the policymaking process differentiates between the budgetary design and the programmes' implementation, it is convenient to model this transformative process using a political economy game between a central authority – the principal – and a collection of functionaries – the agents. We implement this game on a network because policy issues are somehow interconnected. That is, the indicators' values might change not only as a result of public funding but also as a consequence of spillover effects. Network spillovers may affect the probability of increasing or

[1] Note that, while we may use the terms functionary, official, public servant, or bureaucrat for this type of agent, the model makes no assumption on it representing an individual person. Thus, a public-servant agent could also represent an organisation such as a government agency, a ministry, or a technical team, as long as such entity is in charge of implementing the government programmes. This also allows us to account for a larger scope of inefficiency than those that typically arise from individual-level incentives; for example, a poor organisational capacity or a saturated bureaucratic apparatus.

decreasing the value of an indicator if the sum of its incoming links is positive or negative, respectively.[2]

In addition to the network feature, this game takes place in a setting with uncertainty and asymmetric information, so the misalignment of incentives between the principal and the agents in the context of interdependencies produces an allocation problem that is extremely difficult to solve. With this specification, our model accounts for the multidimensional nature of sustainable development. It also characterises governments with diverse aspirations and recognises the existence of a complex world in which policy issues tend to exhibit spillover effects.

For pedagogical reasons, this chapter introduces the model in a top-down fashion. Hence, we start describing the macro-level dynamics and the relevant equations involved. Then, we introduce the political economy game between the government and its officials (or public servants). First, we describe the public servants' decision making in an environment of uncertainty through a reinforcement learning variant. Second, we elaborate on the problem of the government (or central authority) and how we can specify its heuristic strategy. Finally, we provide an overview of the entire structure of the model.

4.1 POLICY INSTRUMENTS

First, let us elaborate further on the concepts of instrumental and collateral indicators introduced in Chapter 3. We begin by describing a country (or subnational entity) with $1, 2, \ldots, N$ policy issues; these could be social, economic, environmental, or technological. In this country, a central authority is responsible for financing government programmes designed to directly improve some of these issues. However, the policy space is, in fact, too diverse and intricate, which makes it impossible for the government to set specific programmes with a mandate to impact each of these issues directly.

[2] Positive spillovers are known in the sustainable development literature as *synergies*, whereas negative spillovers are referred to as *trade-offs*.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

For instance, we could safely assume that not all countries have policy instruments to deal adequately with cybersecurity threats, as this responsibility relies, for the most part, on private companies. Alternatively, other policy issues are too broad or aggregate since their outcomes result from the consequences of many socioeconomic processes. For example, no government in any part of the world has a specific policy programme that directly impacts GDP with a reliable degree of control. Hence, our central authority has direct influence only over $n \leq N$ policy issues, which we say are *instrumental*.

While the government directly affects n indicators, it has goals or aspirations for all the N topics as they are considered socially and environmentally important. Thus, the central authority monitors the country's progress through N development indicators, each describing a high degree of development when presenting a relatively large value. If a policy issue is not associated with a government programme, we say that the corresponding indicator is of the *collateral* type. Accordingly, due to the lack of funding, its progress depends exclusively on spillovers and other exogenous factors that are out of the direct control of the government. Defining indicators as collateral or instrumental depends on the specifics of each country and the government level (e.g., national vs subnational) in charge of decision making. In other words, the model reflects the context that characterises the nature of the country under consideration, which makes the associated inferences reliable for guiding policy prescriptions.

In the next section, we begin explaining how the model operates by describing the mathematical formulation of the indicator dynamics. As the reader will notice, the model is simple enough to avoid unnecessary parameters to calibrate but not too simple to preclude a realistic characterisation of the data in terms of some interesting stylised facts. Our specification establishes, in an intuitive manner, a theoretical framework that describes how the data-generating process might take place. Therefore, in the next step, we develop the micro foundations that provide the theoretical backbone to the main parameters of these equations. We do it by explaining the behavioural model of the public servants and, then, the heuristic approach behind the budgetary allocation of the central authority.

4.2 INDICATOR DYNAMICS

In real-world data, development indicators exhibit various types of behaviour, for example, positive/negative trends, high/low volatility, non-linearities, or periods of continuous/intermittent growth. Statistics-based methods aim at capturing some of these features. A popular modelling framework that allows describing a richer set of these features is stochastic processes; however, often, this type of model lacks a socially relevant theoretical backbone. Our approach in PPI tries to reach a compromise between the flexibility of stochastic processes and the substance of solid theoretical foundations.

From a stochastic-process point of view, we specify a random walk with a drift that allows us to match three critical empirical features: (1) positive and negative trends, (2) initial and final values, and (3) empirical probabilities of growth. As we will show, each feature is associated with a parameter having an intuitive interpretation and an explicit connection to theoretical micro foundations. In other words, the random walk with a drift characterises the macro-level dynamics, while the model's theoretical underpinning lies at the micro level.

The evolution of indicator I_i follows a random walk determined by (1) a probability of growth and (2) a step size. Every period t, indicator i grows with probability $\gamma_{i,t}$, or decreases with probability $1 - \gamma_{i,t}$. If the indicator grows, it does so by a factor (or step size) α_i; if it decreases, it does by α'_i. Let $\xi_{i,t}$ be the binary outcome (0 or 1) of a random draw with probability $\gamma_{i,t}$. Then, we can write the indicator's random walk as

$$I_{i,t+1} = \begin{cases} I_{i,t} + \alpha_i & \text{if} \quad \xi_{i,t} = 1 \\ I_{i,t} - \alpha'_i & \text{otherwise} \end{cases} \tag{4.1}$$

Notice that Equation 4.1 has certain properties that complicate mathematical tractability. First, the probability of success $\gamma_{i,t}$ is not constant. Thus, the probability of going up or down changes over time.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

As we will show ahead $\gamma_{i,t}$ provides a link to the micro foundations by accounting for the dynamic nature of public spending and network spillovers. For now, let us assume that, if indicator i receives more funding or more positive spillovers, then $\gamma_{i,t}$ may increase (and so the chances of the indicator improving).[3]

A way to interpret the binary outcome $\xi_{i,t}$ is in terms of the success or failure of a policy. The more resources or spillovers received, the more likely an existing government programme will succeed. A caveat is in order since such success relates only to short/mid-term funding impacts, not to the possibility of tackling structural transformations. This interpretation holds because we assume, in the simulations, that the existing government programmes remain, for practical purposes, fairly unchanged. Thus, a structurally flawed programme will perform poorly even with substantial financial resources, an idea that we exploit in Chapter 7 to analyse idiosyncratic bottlenecks.

We capture the long-term structural factors through α_i and α'_i. Therefore, if a policy succeeds, the improvement on the indicator is bound by α_i. Notice that α_i is a constant coefficient that does not change over time; hence, it has a long-term nature. When a policy fails, α'_i reflects a cost for the indicator's performance; thus, this parameter also captures structural factors. That is to say, parameters α_i and α'_i account, in a reduced way, for all other factors that contribute to the development of I_i, but that we do not model explicitly. Nevertheless, due to the flexibility of ABMs, it is possible to modify the model to account for additional factors if additional information were available.

[3] It is possible that an indicator value may be unfeasible beyond its theoretical limits (the technical bounds). In this situation, and if the user provides such limits to the model, the simulation conditions Equation 4.1 to not surpass such boundaries. Thus, Equation 4.1 only applies when the realised change does not surpass the technical boundary. Across Monte Carlo simulations, this conditioning induces marginal decreasing returns as, with time, it becomes less likely to reach high(low) indicator values without exceeding the threshold. Again, this is an optional feature that PPI activates only if the user provides information on the technical limits of the indicators.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

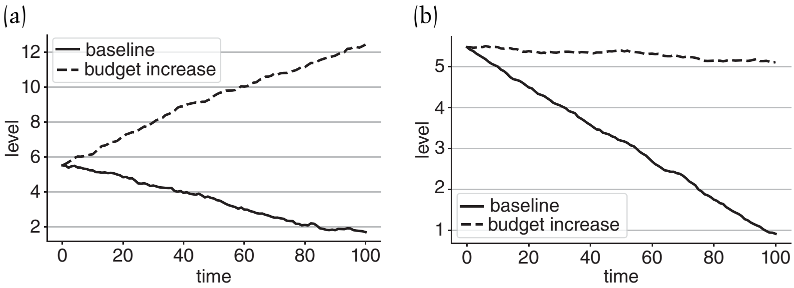

FIGURE 4.1 Illustrative indicator dynamics. Examples of where the budget is (a) effective and (b) ineffective.

Notes: The solid line in Figure 4.1a describes the trajectory of indicator i with the parameter set: $\alpha_i = 0.4$, $\alpha'_i = 0.7$, and $\gamma_i = 0.6$. In Figure 4.1b, the indicator j's trajectory is governed by parameters $\alpha_j = 0.11$, $\alpha'_j = 0.27$, and $\gamma_j = 0.6$. The dotted lines show the dynamics when the probability of success in both indicators increases to $\gamma = 0.7$ by boosting government expenditure.

The specification in Equation 4.1 offers the possibility of modelling the empirical data and generating intuitive counterfactuals that can be informative to policy interventions. Before diving into the model details, it is convenient to provide some examples that clarify how the dynamics behind this equation work. Let us assume that γ_i is constant through time. Then, depending on the values of α_i, α'_i, and γ_i, we can generate different stochastic trajectories. For instance, consider an indicator i described by the parameter set $\alpha_i = 0.4$, $\alpha'_i = 0.7$, and $\gamma_i = 0.6$. Generally speaking, this random walk has a negative drift (due to $\alpha_i < \alpha'_i$), even if the policy is successful most of the time. Next, consider another indicator j with similar dynamics (a negative drift), but governed by parameters $\alpha_j = 0.11$, $\alpha'_j = 0.27$, and $\gamma_j = 0.6$. Figure 4.1 shows the dynamics of both indicators in solid lines (Figures 4.1a and 4.1b, respectively). Each line corresponds to the average trajectory of 100 runs. Both trajectories are qualitatively similar (i.e., they exhibit negative trends), despite coming from random walks with different structural factors.[4]

[4] Note that, together with the parameter values, the difference $\alpha - \alpha'$ is key in determining the trend feature of the indicators. In Chapter 5, we show that, given a potential issue in identifying α_i and α'_i, we can use their difference as the relevant

The dotted lines in Figure 4.1 show the dynamics of policy interventions (counterfactuals), in which the probability of success in both indicators increases from $\gamma = 0.6$ to $\gamma = 0.7$. Suppose that such an increment in the success rate of each policy comes from an increase in public expenditure. As shown by the panels, budget increments are highly effective in indicator i (Figure 4.1a) as they can revert the negative drift into a positive one, despite having $\alpha_i < \alpha_i'$. In Figure 4.1b (indicator j), in contrast, we observe certain expenditure ineffectiveness because budget increments in the same amount as in i are unable to revert the negative trend. Thus, even if there are improvements in the (negative) performance of indicator j, one could say that such a strategy fails when the goal of increasing expenditure is to revert or nullify the negative trend. Next, we elaborate on the micro foundations that generate the probability of success $\gamma_{i,t}$.

4.3 PUBLIC SERVANTS

Let us assume that there exist government programmes for every instrumental policy issue. Furthermore, for each programme, there is a government official (or an agency) in charge of implementing it.[5] Our modelling approach emphasises the transformation of public expenditure into policy outcomes; successful outcomes require efficient use of funding and effective programmes. Thus, the role of each public servant consists of receiving resources from the central authority and transforming them into successful policies.[6] More formally, let $P_{1,t}, \ldots, P_{N,t}$ represent a vector that characterises the distribution of resources across the n instrumental policy issues in period t. By construction, we assume that $P_{i,t} = 0$ all the time if

structural parameter and prove that we can recover its true value from simulated data.

[5] Recall that we have assumed that the government programmes are given and operate without significant changes throughout a simulation.

[6] Even if the allocated spending is used thoroughly in a government programme, the outcome may still be poor as the programme design and operation could be flawed. In this case, we say it needs redesigning due to structural inadequacies (captured in α_i and α_i').

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

indicator i is collateral and, thus, there is no government programme nor associated functionary. We call this vector the *allocation profile* and propose that the central authority determines its configuration. The time sub-index denotes that the government may adjust these allocations dynamically. Something that we explain in more detail in Section 4.4.

Each official i is in charge of their corresponding allocation $P_{i,t}$ (there are $n \leq N$ public servants). While their job is to use $P_{i,t}$ towards the implementation of the corresponding government programme, these public servants may have incentives to be inefficient for several reasons. For example, functionaries may try to promote a future career in the private sector by favouring specific contractors through opaque procurement procedures. Such a strategy may be in the interest of the functionary but not that of the public. Alternatively, the embezzlement of public funds (for the benefit of the servant's pockets or other purposes such as political campaigns) is a well-known case of the inefficiencies that may arise when the incentives of the principal do not align with those of the agent. Such misalignment of incentives gives place to the classic principal–agent problem (Rose-Ackerman, 1975; Klitgaard, 1988); something pervasive in public administrations around the world (in both the Global North and South). Formally, we say that, out of $P_{i,t}$, functionary i uses only $C_{i,t} \leq P_{i,t}$ towards the government programme. Thus, we call $C_{i,t}$ the *contribution* of the functionary and $P_{i,t} - C_{i,t}$ the level of inefficiency.[7]

Public servants face a trade-off when deciding on how many financial resources to contribute to boosting the performance of a policy issue. On the one hand, being a proficient public servant adds to certain political status if such proficiency is reflected adequately in the policy outcomes. On the other hand, a private gain from being inefficient (like the examples mentioned above) may steer the agent

[7] Due to the relative simplicity of the model and the lack of information to calibrate additional parameters, our model cannot disentangle the nature of inefficiencies such as embezzlement, improper public tenders, excessive bureaucracy, and poor administrative practices.

towards a lower contribution. To model this trade-off, let us specify the functionary's benefit function

$$F_{i,t+1} = \underbrace{\Delta I^{*}_{i,t} \frac{C_{i,t}}{P_{i,t}}}_{\text{proficiency}} + \underbrace{(1 - \theta_{i,t}\tau_{i,t}) \frac{(P_{i,t} - C_{i,t})}{P_{i,t}}}_{\text{inefficiency}}. \tag{4.2}$$

Equation 4.2 describes the benefits (or utility) F_i received by functionary i. The equation reflects the previously mentioned trade-off: an addend representing the benefits from proficiency and another one capturing the personal gain from being inefficient. We weight these terms according to the contribution $C_{i,t}$ and the inefficiency $P_{i,t} - C_{i,t}$ as a proportion of the allocation $P_{i,t}$. Proficiency exerts a positive impact on the functionary's political status that, in turn, is signalled by the improvement of indicator $I_{i,t}$ in comparison with the previous period. Such change, however, is relative to the progress made by the other bureaucrats. A relative reference point captures the importance of standing out to gain political status. The relative change of the indicator that brings political status to the government official is

$$\Delta I^{*}_{i,t} = \frac{I_{i,t} - I_{i,t-1}}{\sum_j I_{j,t} - I_{j,t-1}}. \tag{4.3}$$

Coming back to Equation 4.2, in its second addend, we can see that the benefit derived from inefficiency considers $(1 - \theta_{i,t}\tau_{i,t})$, which is a factor related to public procurement. Parameter $\tau_{i,t}$ captures the penalty incurred when an inefficiency is spotted, which we interpret as the *quality of the rule of law*.[8] In its simplest form, $\tau_{i,t}$ could be assumed homogeneous and constant (so $\tau_{i,t} = \tau$ for all indicators and periods), as we do here. Governance indicators could also be considered endogenous variables that change over time for several reasons: a result of government funding, the creation of new laws and regulations, and improvements in the judiciary system.

[8] It is possible to have a governance parameter specific to each indicator, as governance cultures may vary across different sectors or public agencies (such as a permanent ban from holding a public post).

We investigate this modelling choice in Chapter 8. For simplicity, and for most of this book, we assume $\tau_{i,t} = \tau$.[9]

The penalty is only half of the story when speaking of public governance and procurement. The other half is the reduction of opportunities (of being inefficient) through monitoring. We model this utilising a binary variable $\theta_{i,t}$, which takes the value 1 when monitoring has been successful in spotting an inefficiency and zero otherwise. The probability of $\theta_{i,t} = 1$ depends on the *quality of monitoring* and the relative size of the inefficiency. Quality of monitoring, $\varphi_{i,t}$, is another public governance attribute of the country to be analysed. In this case, the institutional variable interacts with the relative size of the inefficiency to produce the probability

$$\lambda_{i,t} = \varphi_{i,t} \frac{P_{i,t} - C_{i,t}}{P_t^*}, \tag{4.4}$$

where $P_t^* = \max(P_{1,t}, \ldots, P_{n,t})$, to make the inefficiencies relative.[10] We use a relative term to measure inefficiencies because the relevance of this feature in a country is often the result of a social convention. For example, even in a society in which corruption is prevalent and, to some degree, accepted, being spotted as a crook or an inept functionary usually requires standing out of the social norm. This social-norm component is an important departure from the traditional economic approach to the principal–agent problem. Thus, our model tries to conciliate the agent-centric economic perspective with the sociological view of social norms and collective action problems (Persson et al., 2013).

Now that we have explained the gains and trade-offs associated with the decisions of public servants, we need to describe the behavioural model through which they determine the contribution levels that they find most beneficial. As in the real world,

9 A more pragmatic reason to assume exogenous governance parameters is that, often, stakeholders have difficulties in finding governance indicators that match the sample period and geographical resolution of their indicator dataset. Thus, to make PPI more accessible and overcome this limitation, we allow for the possibility of setting these factors as exogenous variables.

10 Collateral indicators are not subjected to this process.

functionaries are rationally bounded and face substantial uncertainty due to the complexity of the environment. Hence our modelling choice for the behavioural component is a reinforcement-learning heuristic (Bayer et al., 2009; Carrella, 2014). In particular, we use a specification called *directed learning* in the behavioural literature (and *proportional–integral–derivative controllers* in the engineering literature), in which the bureaucrats either increase or decrease their contributions, depending on the observed change in benefits and the direction of their previous action. The model's intuition is straightforward. If the last action increased the contribution and this produced an increment in benefits, then, in the next period, the public servants will keep increasing their contributions. If, on the contrary, increased contributions generated fewer benefits, then the functionaries will react and change the direction of their actions in an attempt to improve their benefits in the next period.

The actions that functionaries take to increase or decrease their contribution are varied. They depend on the different activities involved in administering government programmes which, in turn, lead to various types of inefficiencies. To abstract from this complexity, we prefer to model all possible actions as a variable $X_{i,t+1}$ that can take any value in the line of the real numbers. An increment $X_{i,t+1} > X_{i,t}$ means that the public servant i increases the size of the contribution, while $X_{i,t+1} < X_{i,t}$ indicates a reduction. For modelling purposes, we are interested in the updating rule of X_i under directed learning. Such a rule can be formalised as

$$X_{i,t+1} = \underbrace{X_{i,t}}_{\text{inertial term}} + \underbrace{\operatorname{sgn}((X_{i,t} - X_{i,t-1})(F_{i,t} - F_{i,t-1}))}_{\text{direction}} \underbrace{|F_{i,t} - F_{i,t-1}|}_{\text{adaptation size}}, \tag{4.5}$$

where $\operatorname{sgn}(\cdot)$ corresponds to the sign function. This operation captures the intuition of directed learning since it reflects the public servants' propensity to maintain the same actions if they are beneficial or reverse them when they do not perform as expected. In contrast, factor $|F_{i,t} - F_{i,t-1}|$ accounts for the size of the change in benefits.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

Once we establish how public servants learn and adapt, the final step is to map their actions into the size of their contribution. We do this through the logistic function

$$C_{i,t} = \frac{P_{i,t}}{1 + e^{-X_{i,t}}}, \tag{4.6}$$

where we specify a positive relationship between actions and contributions. This function guarantees that a contribution cannot be above the resources allocated by the central authority. Hence, the problem of the functionary is to choose a contribution level adaptively with the aim of enhancing their benefit by finding a balance between proficiency and inefficiency (unless extreme governance parameters make an action dominant all the time). Ultimately, decisions depend on the institutional, behavioural, and social mechanisms involved in the policymaking process. Chapter 5 elaborates on the internal validation of this behavioural component.

Before moving on to the next component of the model, let us provide an example of the learning dynamics of the public servants through an illustrative run of the model in Figure 4.2. These plots show the time series of inefficiency (measured as $1 - C_{i,t}/P_{i,t}$) for individual functionaries. All panels use the same parameterisation except for the two governance parameters: φ (for the quality of monitoring) and τ (for the quality of the rule of law). These three examples show (1) how the individual agents respond to different institutional settings and (2) how a social norm of inefficiency emerges.

Figure 4.2a consists of an economy with a good level of public governance in both monitoring and the rule of law (RoL). In this case, agents learn to be proficient, yet sometimes they may attempt to get private gains. The spikes followed by drops in inefficiencies indicate that the punitive mechanisms are effective at correcting this behaviour, as these agents return to the norm relatively quickly. We do not observe this pattern in the scenario described in Figure 4.2b, where monitoring bureaucrats is rare. Since the RoL imposes mediocre punitive measures, agents can afford to venture into higher levels of inefficiency,

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

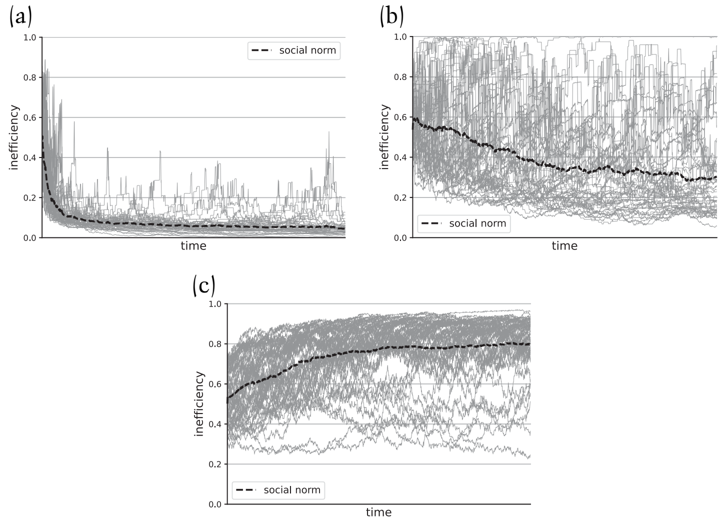

FIGURE 4.2 Learning dynamics and the emergence of social norms under different public governance regimes. (a) Good monitoring and good RoL, (b) poor monitoring and mediocre RoL, and (c) good monitoring and poor RoL.

Notes: We generate these time series from a single run with random parameters and 50 synthetic indicators. Inefficiency is measured as $1 - C_{i,t}/P_{i,t}$. The social norm is the average inefficiency at a given point in time.

as if the penalties were the price to pay for earning a private gain; accordingly, the resulting social norm of inefficiency is higher.

In Figure 4.2c, we present an interesting case that, in our opinion, is appealing for understanding many real-world cases. Here, we have specified a country where monitoring is good, but impunity is widespread. Under this institutional setting, low penalties are a modest price to pay. That is why agents learn it is worthwhile to be inefficient, even when the central authority supervises their actions and catches some of them. This setting corresponds to many mid-income countries with a certain degree of democracy (like some in Latin America). In many of these societies, there is an official discourse and perhaps some actions fostering open governments, transparency, and auditing. Despite the relatively good monitoring mechanisms, when it comes to correcting misbehaviour, it is common to find a poor RoL and prevalence of impunity. In essence, we are in the presence of a government that simulates good governance to improve its public credibility while condoning high levels of inefficiency (and sometimes becoming complicit) through a precarious prosecution.

These simulations show that the model has an explicit and intuitive connection between individual behaviour and social mechanisms. One can explore the various societies that may emerge from different institutional settings using more nuanced governance data, time- and sector-wise. For the time being, it is interesting to see that when we use aggregate and constant empirical indicators as inputs for φ and τ, the emerging levels of inefficiency are consistent with cross-country results from well-known corruption surveys. We show, precisely, this feature as part of our validations in Chapter 5. For now, let us continue explaining the rest of the model.

4.4 CENTRAL AUTHORITY

Now we shift our attention to the government's decision-making process. This player of our political economy game determines the allocation profile $P_{1,t}, \ldots, P_{N,t}$. To speak of allocations, we need first to introduce the total budget B. The modelling of the central authority relies heavily on the availability of expenditure data. That is to say, the granularity of these data establishes the partition level of the budget that the model needs to simulate. To start, let us discuss the ideal case in which there is open spending data for each indicator and period and progressively introduce modelling assumptions as we restrict the granularity of the data. This explanation should provide a comprehensive picture of the flexibility of our framework, as it can accommodate applications with various degrees of data quality and institutional nuances.

Suppose we have highly disaggregated data on how much is spent every period on each government programme. For T periods, this means that we have a *disbursement schedule*

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

$$\mathbb{B} = \begin{bmatrix} P_{1,1} & P_{1,2} & \dots & P_{1,t} & \dots & P_{1,T} \\ P_{2,1} & P_{2,2} & \dots & P_{1,t} & \dots & P_{2,T} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ P_{i,1} & P_{i,2} & \dots & P_{i,t} & \dots & P_{i,T} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ P_{n,1} & P_{n,2} & \dots & P_{n,t} & \dots & P_{n,T} \end{bmatrix}. \tag{4.7}$$

Essentially, the matrix in Equation 4.7 provides a complete mapping between public expenditure and government programmes through time. This setting would be the ideal dataset for undertaking analyses of the relationship between expenditure allocations and indicators' performance. Each row of this matrix is a time series describing budgetary readjustments in each policy issue. If such a dataset exists, our concerns regarding the various motivations of such readjustments would be of second order.[11] Accordingly, the model would take these data as input and simulate the behaviour of functionaries together with the indicator dynamics. Unfortunately, as of today, such a complete mapping does not exist, at least not in a way that allows us to link each expenditure programme to a development indicator without ambiguity. For this reason, we resort to socioeconomic theory in the model and fill these data gaps through simulations.

Next, suppose we have an imperfect mapping between the expenditure partitions and the indicators. One such mapping, for

[11] In our academic work, we try to exploit the temporal resolution of different expenditure datasets to analyse specific periods of potential intervention or the impact of contemporary ephemeral government programmes. To match the temporal coverage of the expenditure data to the indicator data, we produce imputations employing machine learning techniques. Furthermore, because government expenditure typically exhibits temporal trends and the model's parameters are constant through time, an additional treatment to detrend and shift government spending data is required (otherwise, the prospective simulations will be biased. See Guerrero et al. (2023) for an example). This type of analysis is too specialised for the aims and studies covered in this book. Therefore, in all chapters, we preprocess government expenditure by computing the annual average (in each expenditure category) and applying it in each simulated year. This pre-treatment of the data simplifies the analysis substantially and does not affect the results presented in this book in a substantial way.

example, would be expenditure data classified into broad development topics such as national budget tranches or the SDGs. An imperfect expenditure–indicator match introduces uncertainty about the allocation process of expenditure within each tranche (or SDG). $B_{k,t}$ would describe the total expenditure destined to tranche k in period t, and Ω_k a set of policy issues that would potentially benefit from such resources.

To address the uncertainty about the distribution of $B_{k,t}$ across the government programmes in Ω_k, one needs to posess ex ante knowledge about the heuristics that a particular government follows when readjusting its budgets and to be able to incorporate that knowledge into a model. We achieve this using well-known government budgetary heuristics. For example, in the real world, a common practice is to allocate more resources to those policy issues that are relatively laggard, as they may represent potential bottlenecks for development or unsettling problems that require urgent attention due to their political and social repercussions. This is known as 'gapping', and was a promoted practice during the Millennium Development Project. Another criterion relates to public governance and the principle that the central authority tends to reward those agencies (or public servants) who have shown good performances in terms of their indicators. Or, on the contrary, to penalise the inefficient agencies by withdrawing funds.

Of course, getting the heuristic right for a particular country can be challenging. As we have explained in the Introduction, policy prioritisation often seems arbitrary. Nevertheless, our model allows for defining highly specific government behaviour due to the flexibility enabled by its computational nature. Because we apply the model to a large set of countries, as a first approximation, we prefer to implement a heuristic that brings together the principles mentioned above. In addition, we account for the stylised fact of a skewed distribution of budgetary changes that seems prevalent across several countries (Jones et al., 2009) (at least in the punctuated-equilibrium literature from political science). We do this by introducing a stochastic element in the decision rule of the government.

Let $q_{i,t}$ denote the propensity of the central authority to spend in policy issue i during period t. The evolution of the previous propensities is given by

$$q_{i,t} = \underbrace{q_{i,t-1}}_{\text{inertial term}} + \underbrace{U(0,1)}_{\text{stochastic term}} \underbrace{\left\{ \max\left(1, \sum_{h}^{t-1} \theta_{i,h}\right) \right\}^{-1} \sum_{h|\theta_{i,h}=1}^{t-1} \frac{P_{i,h} - C_{i,h}}{P_{i,h}}}_{\text{governance component}}, \tag{4.8}$$

which presents the inertial component of the policymaking process (i.e., political and fiscal lock-ups), and the three heuristics previously mentioned. The variable $q_{i,t-1}$ captures the historical inertia of this propensity. The stochastic component, represented by $U(0,1)$, describes an independent realisation of a random variable with a uniform distribution. The remaining component describes how the current history (until period $t-1$) of spotted inefficiencies in policy issue i is penalised in period t (i.e., for $h \leq t$). The heuristic of prioritising laggard indicators enters through the initial conditions of $q_{i,0}$ in the form of gap closure as

$$q_{i,0} = \max[(G_i - I_{i,0}), 0], \tag{4.9}$$

where G_i corresponds to the development goal of policy issue i. The usability of the gap-closure criterion relies on having information about the goals of each indicator.[12] In contrast, the model sets random initial conditions for the propensities if no such information is available.

A further – optional – refinement requires accounting for structural and idiosyncratic factors that shape the government's expenditure decisions. These factors could be historical, political, or administrative reasons why the government may need to maintain certain

[12] The 2030 Agenda for the SDGs emphasises the setting of precise goals to be reached for each indicator. Likewise, national development plans usually establish quantitative goals for the term a government administration lasts.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

expenditure propensities. Thus, even if there is uncertainty regarding the distribution of $B_{k,t}$ across the indicators in Ω_k, there may be good reasons (or some insightful information) for preserving certain priorities in specific indicators. In Chapter 8, we explore this situation in the context of funding programmes related to public governance. Nevertheless, in this section, we introduce our modelling choice to account for these structural/idiosyncratic factors.

A vector of parameters $b_1, \ldots, b_N$ allows us to account for the structural and idiosyncratic factors of government expenditure.[13] If there is no reason to consider such factors, we assume $b_i = 1$ for every instrumental policy issue. Otherwise, we demonstrate how to exploit them in Chapter 8. For the time being, with these factors, we construct the modulated propensities

$$\dot{q}_{i,t} = \left(\frac{q_{i,t}}{\sum_j q_{j,t}} \right)^{b_i}. \tag{4.10}$$

Finally, to determine the specific allocations that each government programme receives, we employ a function describing the budgetary participation for tranche k.

$$P_{i \in \Omega_k, t} = B_{k,t} \frac{\dot{q}_{i \in \Omega_k, t}}{\sum_j \dot{q}_{j \in \Omega_k, t}}. \tag{4.11}$$

Through Equation 4.11, it is easy to see how we can adapt the model to various data-availability situations. For example, suppose that no data on budgetary tranches exist. In that case, we determine the allocation profile entirely using the total budget such that $P_{i,t} = B_t \dot{q}_{i,t} / \sum_j \dot{q}_{j,t}$. Alternatively, it may be the case that there is no inter-temporal budgetary information. Under these circumstances, we specify the allocation profile by $P_{i,t} = B \dot{q}_{i,t} / \sum_j \dot{q}_{j,t}$, where B is an expenditure amount that does not change across time. Furthermore, it could be the case that no budgetary information exists. Then, we can simply set B to an arbitrary number such as 1, and perform counterfactual analysis in relation to that baseline. Of course, the

[13] Once more, collateral indicators do not go through this process.

less budgetary data available, the weaker the inferences we can obtain and the more limited our interpretations and comparisons can be. Nevertheless, having the capability of producing counterfactuals in a model with an explicit expenditure–development connection is still an advantage over alternative frameworks. Fortunately, the data revolution and the Open Gov movement have contributed, in recent years, to the availability of this type of information, so this book will show applications that use budgetary data with various degrees of granularity.

Before proceeding to the next section, we should highlight that the model can accommodate other budgetary nuances. For example, it could be the case that one government programme receives funds from multiple tranches. For instance, in the SDG official dataset (United Nations, 2020), some indicators are classified into multiple SDGs.[14] Computationally speaking, accommodating this feature is trivial. Mathematically, it only requires Equation 4.11 to be changed into

$$P_{i \in \Omega_k, t} = B_{k,t} \frac{a_{i,k} \dot{q}_{i \in \Omega_k, t}}{\sum_j a_{j,k} \dot{q}_{j \in \Omega_k, t}}, \tag{4.12}$$

where $a_{i,k}$ indicates the weight that tranche k represents in the budget of programme i. Of course, this requires prior knowledge about the relative proximity of government programmes – and their expenditure – to the indicators in the different tranches or SDGs. At the cost of double-counting expenditure, for simplicity, we prefer to assume $a_{i,k} = 1$ given the lack of information to establish the alternative specification. As shown in the next section, this double counting is not a problem since the model's free parameters counterbalance this effect after we calibrate them.

4.5 DEVELOPMENT OUTCOMES

Finally, let us explain the last piece of the model: the one that connects the micro-level mechanisms to the macro-level dynamics:

[14] Similar to indicators 7.b.1 and 12.a.1 and indicators 4.7.1, 12.8.1, and 13.3.1 among others.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

the probability of success $\gamma_{i,t}$. This probability is an endogenous dynamic variable that depends on (1) the resources channelled by the government programmes and (2) the spillovers received from other policy issues, either instrumental or collateral. Hence, the probability of success of government programme i is determined by

$$\gamma_{i,t} = \underbrace{\beta_i}_{\text{expenditure returns}} \times \left[\underbrace{C_{i,t}}_{\text{contribution}} + \underbrace{\frac{1}{1+B_t}\sum_{j}^{N} C_{j,t}}_{\text{systemic efficiency}} \right] \times \underbrace{\left(1+e^{-S_{i,t}}\right)^{-1}}_{\text{spillovers}}. \tag{4.13}$$

The probability of success of a government programme depends, first, on the contribution of the public servant in charge. Therefore, if the public governance of an economy is poor, the impact of an allocation $P_{i,t}$ will be limited because the functionary will have incentives to set low contribution levels. Notice that, in the case of a collateral indicator, $C_{i,t} = 0$ all the time, systemic efficiency (the ratio of total contributions to the budget) is also relevant for the probability of success, as it connects the performance of the indicators to the general "financial health" of the system. The latter also affects collateral indicators since government expenditure is an important component of aggregate demand and promotes general development through crowding-in factors.[15]

To normalise the budgetary information in such a way that $\gamma_{i,t} \in [0,1]$, we introduce the free parameter β_i. One way to interpret these parameters is to conceive them, partially, as the returns to the expenditure in the corresponding instrumental indicators. For example, suppose that two indicators, i and j, exhibit the same performance, but i receives 100 times more resources than j. When calibrating their respective βs, we would obtain that β_j is approximately two orders of magnitude larger than β_i. Assuming all else constant, one would interpret that each dollar spent in β_j pays back 100 times more than

[15] The number one in the denominator $1 + B_t$ makes sure that the efficiency term is defined, even under the complete absence of public funds.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

in β_i. Such differences are common in real-world policymaking, as there exist topics that inherently require substantially more resources than others to exhibit similar proportional improvements in the associated indicators (e.g., government programmes related to public infrastructure).

Finally, the remaining factor captures the spillover effects $\left(1 + e^{-S_{i,t}}\right)^{-1}$, where $S_{i,t}$ corresponds to the total amount of spillovers (positive and negative) that policy issue i receives in period t. We compute these incoming spillovers every period using the adjacency matrix $\mathbb{A}$ (of size $N \times N$), where entry $\mathbb{A}_{j,i}$ denotes the spillovers that, if realised, policy issue j sends to i. Thus, $S_{i,t} = \sum_j \mathbf{1}_{j,t}\mathbb{A}_{j,i}$, and $\mathbf{1}_{j,t}$ is the indicator function: 1 if indicator j grew in the previous period and zero otherwise.[16]

As we have discussed in Chapter 3, the network in $\mathbb{A}$ does not represent causal relations between the indicators. Instead, it captures structural interdependencies between the development indicators. By incorporating $\mathbb{A}$ as an exogenous variable in the model, we can include specific information concerning each country's internal structure (its context). Notice that the spillovers are short-term events that realise from the growth dynamics of other indicators. Hence, and to be consistent with our previous argument, the network reflects the long-term conditional dependencies that influence the realisation of the spillovers. Empirically speaking, the matrix $\mathbb{A}$ is an input of the model, and we can construct it in various ways (quantitative and qualitative). We elaborate on our method of choice for estimating $\mathbb{A}$ in Chapter 5.

[16] Note that we consider only spillovers (positive and negative) generated from growth in indicators, not declines. This is mainly because most of the discussions around synergies and trade-offs have to do with the effects of improving certain policy issues. While there are also spillovers from worsening dynamics, they are ill defined when thinking of a positive spillover when a policy issue worsens (with the exception of environmental improvements from a decline of industrial expansion). Thus, we decided that, for this version of the model, we need to focus only on the externalities produced by improvements, as policies usually aim at producing positive outcomes.

Table 4.1 *Variables of the model*

Symbol	Variable	Source	Type
α_i	limiting structural factors	calibrated	exogenous
α'_i	structural costs	calibrated	
β_i	expenditure returns	calibrated	
b_i	expenditure modulation parameter	imputed or calibrated*	
$I_{i,0}$	indicator initial level	indicators' time series	
$I_{i,-1}$	indicator final level	indicators' time series	
$\mathbb{B}$	disbursement schedule	open-spending data*	
$\mathbb{A}$	spillover network adjacency matrix	indicators' time series*	
$\varphi_{i,t}$	quality of monitoring	worldwide governance indicators**	
$\tau_{i,t}$	quality of the rule of law	worldwide governance indicators**	
$G_{i,t}$	development goal	development plans/documents*	
$X_{i,t}$	agent actions	Equation 4.5	endogenous
$F_{i,t}$	agent benefits	Equation 4.2	
$C_{i,t}$	agent contributions	Equation 4.2	
$\lambda_{i,t}$	probability of spotting inefficiencies	Equation 4.4	
$\theta_{i,t}$	binary outcome of monitoring	Equation 4.2 & Equation 4.4	
$\gamma_{i,t}$	probability of successful growth	Equation 4.13	
$\xi(\gamma_{i,t})$	binary outcome of random growth process	Equation 4.13 & Equation 4.1	
$q_{i,t}$	propensity to spend	Equation 4.8	
$q_{i,0}$	initial propensity to spend	Equation 4.9	
$\dot{q}_{i,t}$	modulated propensity to spend	Equation 4.10	
$S_{i,t}$	net incoming spillovers	Equation 4.13	
$P_{i,t}$	government allocation	Equation 4.11	
$I_{i,t}$	indicator level	Equation 4.1	

Notes: * data on these parameters are optional. ** these parameters can also be endogenous if their respective indicators are part of the policy space on which the government allocates resources.

4.6 SUMMARY AND CONCLUSIONS

As a summary, Table 4.1 presents all the variables and parameters of the model. Note that only three free parameters are employed: α, α', and β, and that they help to describe meso and macro considerations about the relationship between public expenditures and indicator performance. The micro-level variables, in contrast, are all endogenous. This feature of the model is appealing since we do not need to worry about finding micro-level data to parameterise functional forms. In this manner, we avoid a major limitation of many behaviourally driven ABMs. In Chapter 5 we show that the endogeneity of the behavioural mechanisms is not trivial since it is responsible for the emergence of intuitive and stable results.

In Figure 4.3, we provide a diagrammatic depiction of the model's transformation mechanisms, starting at the bottom with the

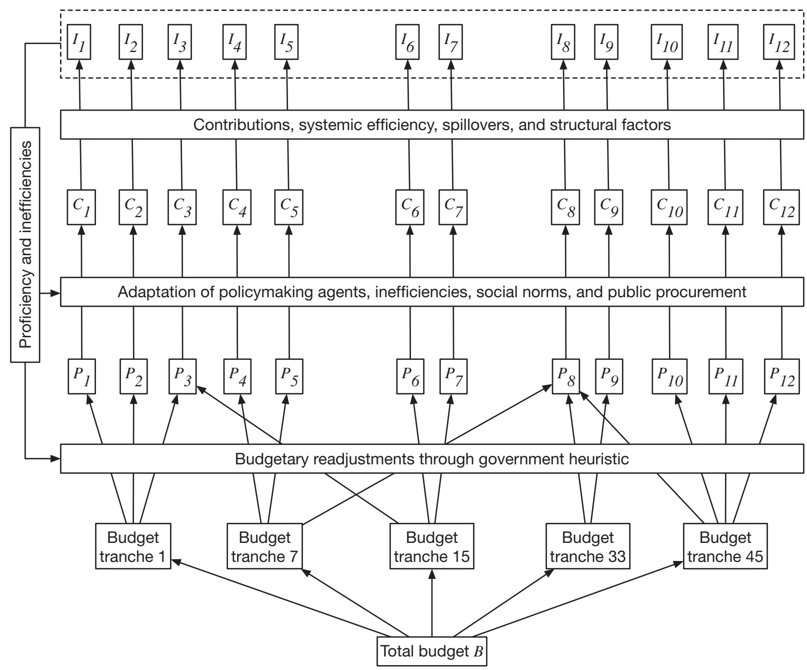

FIGURE 4.3 Bottom-up and top-down structure linking government expenditure and indicators.

https://doi.org/10.1017/9781009022910.007 Published online by Cambridge University Press

expenditure allocations, and ending at the top with the indicator dynamics. First, once the model allocates the budget endogenously, the indicators' dynamics ensue in a decentralised yet interactive

Algorithm 1 Model pseudocode

```
foreach period t do
    foreach public servant i do
        receive public funds P_{i,t};
        evaluate the benefits from the previous contribution
        C_{i,t-1};
        establish new contribution level C_{i,t};
    foreach indicator i do
        if the indicator is instrumental then
            implement public policy using the resources C_{i,t};
        receive the incoming spillovers S_{i,t};
        determine the probability of success γ_{i,t} according to C_{i,t}
        and S_{i,t};
        if the public policy is successful (with probability γ_{i,t})
        then
            improve the indicator according to the long-term
            structural factors α_i;
        else
            worsen the indicator according to the long-term
            structural costs α'_i;
    the government monitors the policymakers through
    imperfect mechanisms;
    the government penalises those who are found being
    inefficient;
    the policymakers receive the benefit from their chosen
    contributions;
    the government updates the allocation profile P_{1,t}, ..., P_{n,t};
```

fashion. This process fundamentally differs from traditional economic models in which homogeneous agents respond to a centralised set of incentives (e.g., a price vector), avoiding interactions or sociological considerations such as social norms. We can see how different behavioural and network elements contribute to the indicators through the model's bottom-up flow. Then, these aggregate dynamics feedback from top to bottom, generating multi-scale feedback loops. One of the virtues of specifying agent-level behaviour is that we can make inferences at both the macro and micro level for validating the model in ways that are not possible with more traditional approaches. Such validation strengthens the confidence in inferences derived from experiments of policy interventions using counterfactual simulations. This diversity of validation tests is possible because the model includes nuances related to the adaptation of the agents to interventions and systemic effects that one cannot capture with models specified exclusively at the micro level.

Finally, in Algorithm 1, the reader can see the procedural logic of the model in the form of pseudocode. This set of instructions emphasises the timing of the agents' decisions (central authority and public servants) and the micro–macro link connecting the budgetary decisions with the indicators dynamic. From this algorithm, it is clear that government expenditure is an endogenous variable whose allocations are influenced by the public servants' decisions to be proficient or inefficient when receiving budgetary funds.