7 Government Spending and Structural Bottlenecks

When performing quantitative evaluations on how budgetary changes impact development indicators, it is not enough to estimate the sensitivity of these indicators through econometric or machine learning models. While these types of statistical analysis are informative about the response that policy issues had or may have to a reallocation of resources, they also miss critical considerations regarding the dynamics of the indicators and the short-term limits of government spending. One such consideration is that the pattern of disbursements can produce specific outcomes. Another is the presence of *structural bottlenecks* that constrain the indicators' evolution. Informally, a bottleneck is a set of factors precluding the possibility that some variables grow as expected. The concept of structural bottleneck has been treated differently across the development literature.

In some settings, structural bottlenecks are described as an endemic problem of limited inputs or missing infrastructure that hamper a country's development potential. In others, they are conceived as a disruption in the supply of critical resources (e.g., raw material, energy, intermediate, and capital goods) or in the functioning of the supply chain, firms' logistics, or the freight transport system. In any case, the concept of structural bottlenecks remains loosely defined. Consequently, there are not many quantitative tools to analyse it. In this chapter, we investigate the idea of structural bottlenecks and provide a first formalisation through PPI; later, in Chapter 11, we develop a second formalisation. Hence, we aim to provide a comprehensive framework to understand and quantify structural bottlenecks in a setting of multidimensional sustainable development.

7.1 ON THE CONCEPT OF A STRUCTURAL BOTTLENECK

In the context of PPI, we formalise the idea of structural bottlenecks when analysing short-term impacts on indicators. Furthermore, we identify two types of structural bottlenecks relevant to the studies conducted in this book: *idiosyncratic* and *systemic* bottlenecks. We focus, in this chapter, on idiosyncratic bottlenecks while reviewing the systemic bottlenecks in Chapter 11. An idiosyncratic bottleneck arises when structural factors prevent the development of a policy issue, despite public expenditure being readily available. To illustrate this idea, imagine a public agency with a portfolio of government programmes, each designed to impact a different policy issue. Some of these programmes could have an excellent design, and their implementation may enjoy a well-trained bureaucracy.

For example, in a programme for incentivising green-technology investments, it may be that the relevant public servants have the right technical skills to understand the potential of different technologies. Alternatively, it would be the case that the ministry in charge of this programme has, internally, excellent mechanisms of governance, minimising collective action problems that could be conducive to the misallocation of resources. Increasing the expenditure on such a programme can lead to a significant impact. In other words, one would expect a relatively sensitive (positive) response of the associated indicators to higher expenditure in this programme. But one would expect the opposite outcome in the case of a poorly designed green technology programme whose operation includes ill-conceived public tenders and inadequate psychical or technological infrastructures. In other words, if something structural to the programme or the environment in which it performs is unfit for purpose, one should expect a poor response or insensitivity from the relevant indicators, even when the programme is well-funded.

As explained in earlier chapters, reforming a government programme involves several tasks, such as investing in infrastructure, developing public servants' human capital, and creating auditing processes. Typically, the implementation of such tasks is expected to

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

materialise in the long term, so these issues are considered structural. From a short-term perspective, in contrast, the type of decisions that can impact an indicator relate, essentially, to how many resources are involved and how to allocate them to the existing programmes. Such distinction between long and short-term tasks is not explicit in more traditional evaluation frameworks. In PPI, we represent this distinction by invoking different parameters: the success probability γ_i and the difference $\alpha_i - \alpha_i'$. That is to say, in the short term, a government may be able to increase the probability of success γ_i of an existing programme by enlarging the sum of resources it receives and spends. However, the value of the structural difference $\alpha_i - \alpha_i'$ constrains the progress achieved through more spending. To address the bounds imposed by $\alpha_i - \alpha_i'$, one would need to implement long-term structural changes such as those mentioned above.

7.1.1 A Formalisation: The Budgetary Frontier

In PPI, we formalise the idea of an idiosyncratic bottleneck when thinking in a hypothetical situation where a government has all the necessary resources to guarantee the success of its existing programmes.[1] For our model, it implies that the probability of success γ_i of an instrumental indicator i would always have a value of 1. That is to say, irrespective of the changes in the model's endogenous variables, we produce our simulations by setting $\gamma_{i,t} = 1$, counterfactually, in each instrumental indicator. We call this scenario the *budgetary frontier*, a concept we introduced in Guerrero and Castañeda (2022). While the budgetary frontier is an unrealistic hypothetical scenario, it offers an insightful benchmark to unearth information regarding structural bottlenecks not evident under methodologies that do not distinguish between short-term and long-term government policies.

To illustrate the usefulness of the budgetary frontier, suppose that an anti-poverty programme succeeds in half of the years it operates (so $\gamma_{i,t} = 0.5$ on average) and that, under its current funding level, it is expected to reach its goal of poverty eradication in 20 years.

[1] Even to overcome side effects like those created by incoming negative spillovers.

Next, assume that we simulate this programme while operating at the frontier and observe that the new expected time to eradicate poverty is 19 years. Here, even if public expenditure has increased enough to guarantee a success rate of 100%, the impact on poverty is negligible, as we are shortening the poverty-eradication horizon by only 1 year. This scenario shows a situation in which, despite abundant financial resources, the structural parameters α_i and α_i' severely constrain the improvement of the indicator (see Figure 7.1 for a graphical description of bottlenecks).

Recall that we calibrate the structural parameters using historical data. Thus, this outcome means that, historically, there

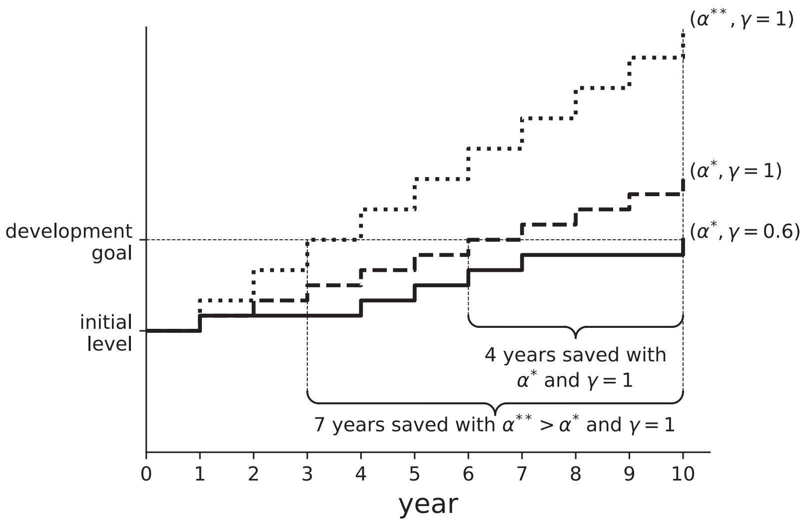


FIGURE 7.1 Illustration of idiosyncratic bottlenecks as a result of structural constraints.

Notes: This chart illustrates the different dynamics of a hypothetical indicator assuming that α_i can take two values: α^* or α^{**} such that $\alpha^* < \alpha^{**}$. We also assume $\alpha_i' = 0$ and $\gamma_{i,t} = \gamma$ where γ can be 1 or 0.6. The inequality $\alpha^* < \alpha^{**}$ means that an indicator where $\alpha_i = \alpha^*$ is subjected to more stringent structural constraints when compared to that with $\alpha_i = \alpha^{**}$. In addition, $\gamma = 1$ means that the indicator operates on the budgetary frontier. The dashed line shows a marginal improvement over the baseline case (solid line) through an increase in financial resources (a short-term intervention). However, it remains restricted by a structural bottleneck. In contrast, the dotted line shows a substantial improvement in performance resulting from more resources and a relaxation of the structural bottleneck.

are problems with the anti-poverty program, which are unrelated to short-term financial considerations. In this example, the budgetary frontier has helped identify a potential bottleneck when the years saved are relatively few, despite assuming enough public funds. Although the model does not identify the factors behind this structural inefficacy (e.g., lack of capacity, poor infrastructure, or inadequate tenders), it allows the detection of policy issues experiencing idiosyncratic bottlenecks. This information is very valuable as part of any diagnostic framework that aims at operating in a multidimensional setting. Besides, identifying all potential structural factors across all these dimensions would be impossible when using only the data we employ in PPI.

7.2 SIMULATION STRATEGY

Our simulation strategy is straightforward. We deploy the same data and calibration used in Chapter 6, and focus on generating counterfactuals (for each country) on the budgetary frontier. Once the baseline and counterfactual outputs are generated, we compare their resulting gaps and assess how sensitive are the different indicators to operating on the frontier. Then, we combine this information with the historical performance of the indicators to develop a methodology that allows for identifying potential idiosyncratic bottlenecks. Finally, we develop a flagging system to differentiate between idiosyncratic bottlenecks according to the 'urgency' to unblock them. We discuss the details of each strategy as we present our results in the next section.

7.3 RESULTS

7.3.1 *Insensitivity on the Budgetary Frontier*

First, let us look at the difference between the estimated gap in 2030 (the benchmark) and the gap expected if countries would operate on the budgetary frontier. In Figure 7.2, we show these differences for each indicator and country. We present our estimates in separate scatter plots according to the corresponding country group. The horizontal axis presents the benchmark expected gap, while the vertical

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

shows the expected gap prevailing on the frontier. Since, by construction, the budgetary frontier eliminates negative trends, the expectation is that a frontier gap cannot be larger than the benchmark gap, so every dot in the plot should be no higher than the 45-degree line. Furthermore, if indicators were sensitive to budgetary increments, all the dots would lie far away from the diagonal. In an ideal scenario, an indicator would lie on the horizontal axis as it would mean that, on the frontier, the corresponding development gap would be closed.

As we can see from Figure 7.2, the ideal scenario of a complete gap closure under the frontier does not occur frequently. Nonetheless, many indicators (dots) exhibit a substantial reduction in expected gaps when operating at the budgetary frontier. Sometimes, the reduction is higher than 10%, as indicated by the SDGs' average of each country group (see inset legends). Notice that, for the West countries, the average reductions are generally smaller than in countries of the other groups. In contrast, it is more common to find in Africa substantial reductions in gaps when operating at the frontier. These outcomes relate to the problem of diminishing marginal returns, indicating that development gaps in the Global North are relatively small, so it is more difficult to close them, even when operating at the budgetary frontier.

Likewise, the colours of the dots do not seem to be distributed randomly within the different charts. This outcome suggests that the degree of sensitivity depends, among other things, on the specific SDGs and indicators intervened (i.e., there exist uneven impacts across policy issues). From these charts alone, one cannot infer the existence of bottlenecks. To detect bottlenecks, we require additional information on the historical performance of the indicators. Presumably, an idiosyncratic bottleneck needs to meet two conditions: (1) a poor performance (historically) and (2) a low sensitivity of the associated indicators to budgetary changes. These conditions imply that, when the gaps are already small, insensitivity to budgetary changes may not be a warning signal about something structurally wrong with the associated programmes.

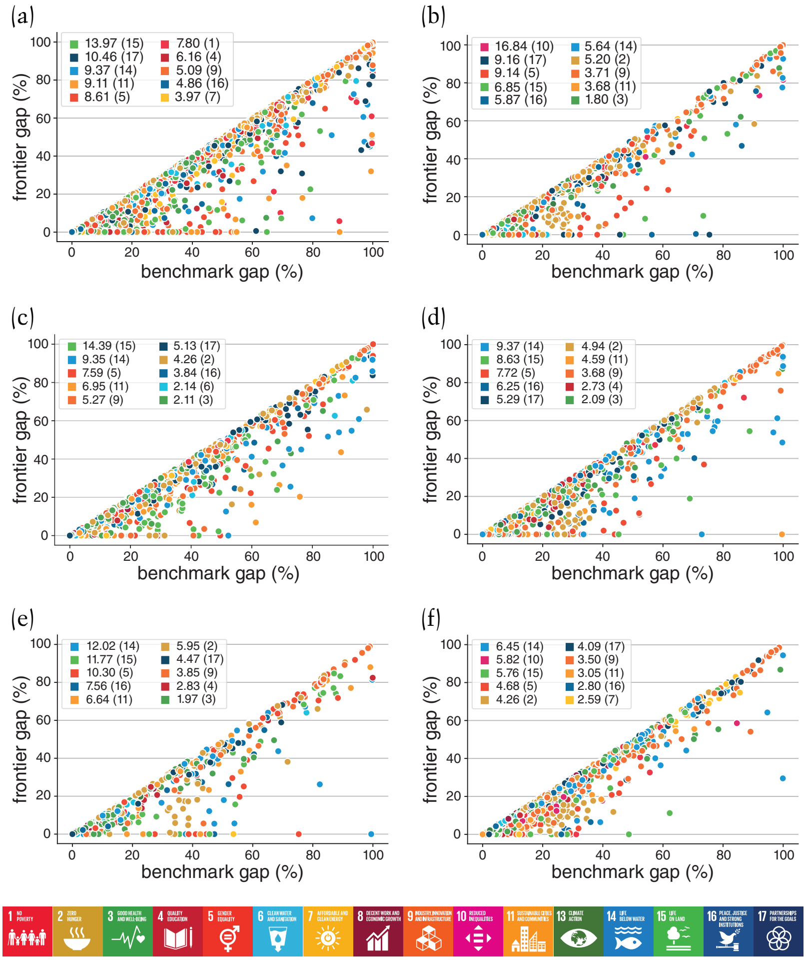


FIGURE 7.2 Reductions in expected gaps under the budgetary frontier. (a) Africa, (b) Eastern Europe and Central Asia, (c) East and South Asia, (d) LAC, (e) MENA, and (f) West.

Notes: Each dot corresponds to an indicator in a specific country. The inset legend shows the ten SDGs that exhibit the largest average reductions in gaps within a group of countries. We compute the size of the reductions as the difference between the benchmark gap and the frontier gap. The average reduction is the number next to the coloured marker (in per cent units), and the corresponding SDG is indicated in parenthesis.

Sources: Authors' calculations with data from the 2021 Sustainable Development Report.

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

7.3.2 *Poor Performance + Insensitivity = Bottleneck*

Insensitivity to budgetary increments does not necessarily mean that a government programme is ill-conceived. Low sensitivity could emerge because of a heterogeneous population and the presence of marginal inhabitants with special requirements. This explanation coincides with the economic concept of decreasing marginal returns. For instance, once most of the child population is fully vaccinated, it becomes extremely difficult to protect more children, perhaps because their parents have religious beliefs against vaccines. In another example, water access and sanitation can improve swiftly when a government establishes adequate infrastructure around a metropolitan area with a large population. Yet further improvements are cumbersome when people in isolated areas are not covered. In fact, in various countries, policies promoting migration to more populated communities have been devised precisely to overcome the constraints imposed by decreasing marginal returns to public expenditure and by the insurmountable challenges (technical, financial, or time-wise) associated with implementing structural changes.

For identifying structural bottlenecks, it is necessary to consider not only a potential gap reduction but also the historical level of an indicator. The exact definition of what it means to have a low performance involves a certain degree of subjectivity, as researchers or policymakers can establish arbitrary thresholds. Hence, the use of expert knowledge to determine appropriate levels is recommendable. We do not intend to establish a thoroughly thought criterion for separating good from bad performers. Our aim is only to show how to deploy PPI for determining idiosyncratic bottlenecks once thresholds are chosen. With that aim, let us begin by characterising the space defined by the gap reductions and the historical performance of the indicators. We do so by counting the number of indicators (across countries in the same group) that belong to a discrete combination (a 2D bin) of average indicator level (inter-temporal mean) and gap reduction (under the budgetary frontier).

On the left horizontal axis, we show the indicators' performance. On the right, we express the difference between the baseline and the frontier gaps, measured as a percentage of the baseline. We colour the 3D bars of the histograms according to different thresholds along both horizontal axes so that the warmest tones denote the worst outcomes in both dimensions. For example, if an indicator's historical level is between the 0–20 bin, and a reduction of more than 90% is expected in its gap (under the budgetary frontier), then it appears in the frontal corner of the plot with a colder tone. The height of the bar in the (0–20)-to-(90–100) coordinate indicates the number of indicators with those performances and sensitivities. The plots show these counts in logarithmic scales because most indicators fall into the back corner.

The intriguing cases in these plots appear in the coordinates that refer to low levels of performance and sensitivity (the red 3D bars located in the right corner of each panel). Arguably, if there are structural bottlenecks, they must be in this region because low-level indicators would have to close relatively large gaps. Thus, if they also present low sensitivity to the budgetary frontier, one can assert that such insensitivity is not due to decreasing marginal returns (or population heterogeneity). The latter outcome occurs, presumably, because long-term constraints may be dampening the potential impact of public expenditure.

For the sake of illustration, let us determine an arbitrary threshold of 50% to separate good and bad historical performance, as well as low from high gap reductions. Then, we colour in blue all bars outside the area defined by this threshold and use warm tones for those inside. The first thing to notice in Figure 7.3 is a flat triangular area of blue. The absence of bars in this area suggests a lack of indicators, around the world, performing well, with promising perspectives, or meeting both conditions. The flat area is relatively large in the groups of LAC and MENA. Second, the non-flat segment of the blue area indicates the presence of several indicators whose coordinates are below one of the two 50% thresholds. These cases tend to have poor performance

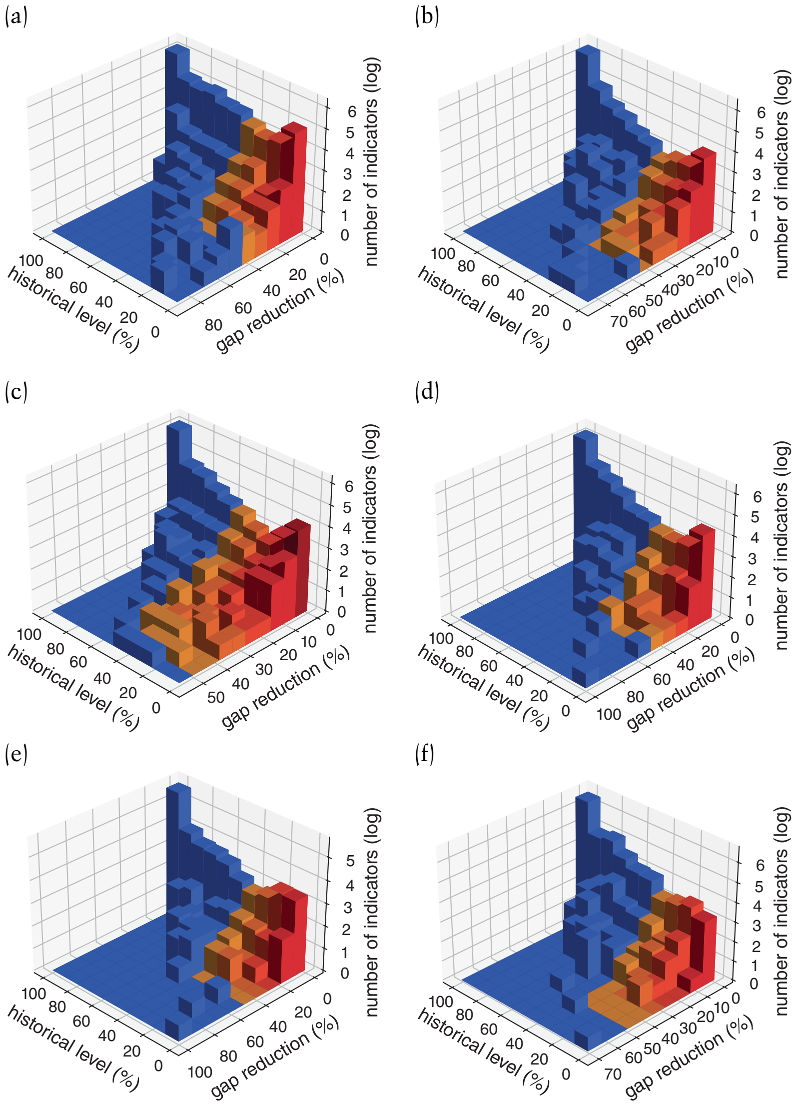


FIGURE 7.3 Distribution of indicators according to historical performance and potential gap reductions. (a) Africa, (b) Eastern Europe and Central Asia, (c) East and South Asia, (d) LAC, (e) MENA, and (f) West.

Notes: Histograms of development indicators using historical levels and gap reductions. We measure historical levels as the inter-temporal average level of each indicator (from the sample period; normalised and expressed in percentages). We compute the size of the reductions as the difference between the baseline gap and that calculated on the budgetary frontier. The indicator counts in the vertical axis are on a logarithmic scale. We make these counts on indicator-country combinations.

Sources: Authors' calculations with data from the 2021 Sustainable Development Report.

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

or an insignificant sensitivity but not both. Under our definition, these indicators cannot be considered bottlenecks. Notice that the largest count of indicators with low performance but high sensitivity corresponds to Africa, while the smallest counts are in the West and East and South Asia. Finally, the number of potential bottlenecks identified with the warmest tones (those closest to the right corner of each panel) is significant in the six groups of countries, yet the case of East and South Asia is salient.

Next, let us isolate all those indicators that fall within the red areas of Figure 7.3 to study where the potential bottlenecks situate across the SDGs. Figure 7.4 shows these distributions through treemap plots for each country group. When an SDG does not appear in a chart, it implies no evidence of potential idiosyncratic bottlenecks in the corresponding group. Nevertheless, the reader should be careful when interpreting a lack of evidence related to bottlenecks, as some SDGs have poorer data coverage when compared to others.[2]

Notice that, across all groups in our dataset, the bottlenecks in SDG 9 dominate, especially in Eastern Europe and Central Asia. Interestingly, the prevalence of SDG 9 as the largest category of potential bottlenecks is not a consequence of the number of indicators. As suggested in Table 3.1, there are other SDGs with more indicators than SDG 9 (e.g., SDGs 2 and 3).[3] In contrast, we do not find potential bottlenecks in SDG 8 in any of the groups. We identify a few bottlenecks in SDG 4 in LAC and Africa. Although poverty is a prevalent issue in LAC, the simulation results do not indicate the presence of structural constraints in the associated government programmes. Hence, the poor performance of the associated indicators seems caused by a lack of funding. This result suggests an opportunity space for these countries by generating short-term impacts through

[2] Recall that, in the SDR dataset, there are no observations for SDG 12 that are usable under PPI in any country. In addition, SDG 10 is quantified nowadays in just a handful of countries.

[3] When counting all indicators across countries, SDG 9 is not even fifth by most counts.

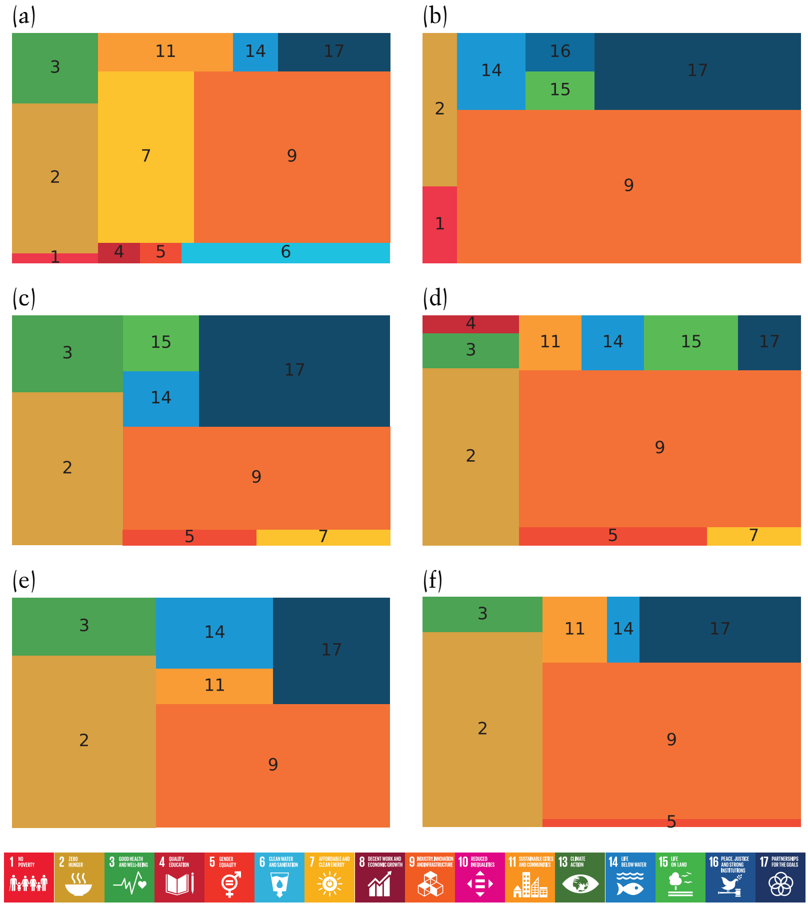


FIGURE 7.4 Distribution of structural bottlenecks across SDGs by country groups. (a) Africa, (b) Eastern Europe and Central Asia, (c) East and South Asia, (d) LAC, (e) MENA, and (f) West.

Notes: Each treemap plot shows the distribution of structural bottlenecks across SDGs within each country group. We perform these counts on indicators identified below 50% in historical levels and sensitivity to the budgetary frontier, i.e., those located in the coordinates with red bars in Figure 7.3. We make these counts on indicator-country combinations.

Sources: Authors' calculations with data from the 2021 Sustainable Development Report.

a shift in policy priorities (i.e., more funding) towards anti-poverty government programmes.

7.3.3 Not All Bottlenecks Are Created Equal: A Flagging System

While identifying structural bottlenecks can be extremely helpful when allocating budgetary resources, a government may still face stringent constraints that prevent it from attending to all of these issues in a single administration. In this case, it is indispensable to differentiate between bottlenecks and identify those that are 'less desirable' than others. For this, we develop a flagging system that categorises indicators with idiosyncratic bottlenecks in three types (or flags): yellow, orange, and red. To explain the proposed classification scheme, let us refer to Figure 4.1. We show, in this figure, two examples of indicators with a projected negative trend. In the first case, the trend is reverted by increasing public spending. In the second case, government expenditure improves the indicator, but it cannot reverse the negative trend.

Let us use these illustrations to motivate our flagging system, assuming that the intervention consists of operating at the budgetary frontier (so $\gamma_{i,t} = 1$). Our first case consists of indicators with a positive trend. Here, an idiosyncratic bottleneck means that structural factors slow down the improvement of the indicator, to the point of not being able to close its development gap, even with substantial resources. While we interpret this as a bottleneck, we do not expect the indicator to worsen. Thus, if policymakers would conduct business as usual, this policy issue would not deteriorate. We categorise these indicators under the yellow flag.

The second case corresponds to that illustrated in Figure 4.1a. Here, the indicator projects a negative trend, but policymakers could reverse this tendency with enough expenditure. Since we are dealing with a bottleneck, this reversion would not be enough, generally, to close a substantial portion of the development gap. Nevertheless, it would at least stop the indicator from worsening. We classify these

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

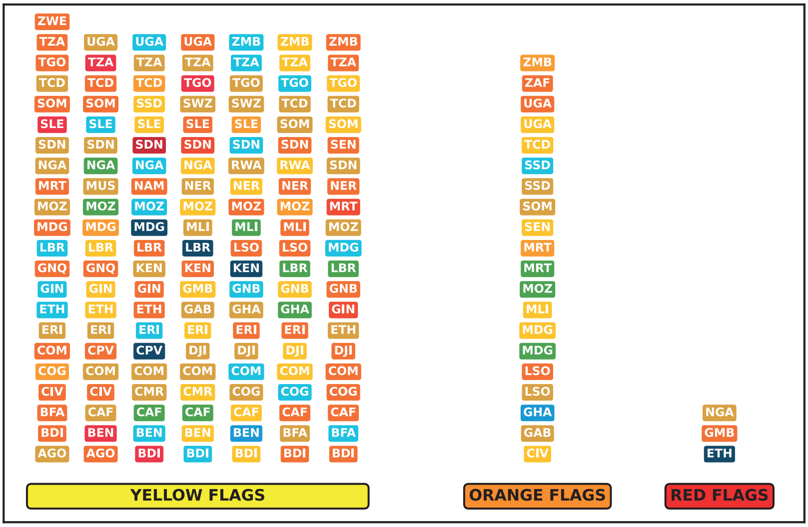
cases under the orange flag. Finally, we depict the third case in Figure 4.1b. Here, the indicator shows a negative trend, and public expenditure cannot revert it. Thus, we employ a red flag because policymakers cannot stop the worsening of this policy issue; not even with substantial financial resources. This type of bottleneck needs the urgent attention of policymakers.

A meticulous reader would have noticed by now that, under the budgetary frontier, red flags would not be detectable since the assumption of $\gamma_{i,t} = 1$ guarantees a positive trend. Nevertheless, for practical purposes, we can think of a *relaxed budgetary frontier*, in which $\gamma_{i,t}$ is close to 1. For example, suppose $\gamma_{i,t} = 0.9$, which means that, in 9 out of 10 times, the policy is successful. Under a relaxed frontier, we are in the presence of red flags when bottleneck indicators cannot revert their negative trend. Of course, this new threshed is also somehow arbitrary. Nevertheless, being able to discriminate between potential idiosyncratic bottlenecks, even when using a subjective threshold, is a valuable task to understand structural constraints in sustainable development.

Using the 90% relaxed frontier, we classify all bottlenecks in the SDR dataset according to our flagging system. We present the results in Figure 7.5 by country groups. Each marker denotes an indicator in a specific country (the name of the country is inside the marker). We colour the markers according to the SDG to which the indicators belong. The most salient result is that, in all groups, the red flags are the least common (even more, East and South Asia and MENA have none), followed by the orange flags. However, in relative terms, orange flags are not frequent in the West and Africa, while red and orange flags are common in Eastern Europe and Central Asia. Governments should be very cautious when funding policies of indicators with orange and red flags, and pay more attention to revising the inner works of their associated programmes.

Importantly, the bottleneck identification and classification procedure proposed in this chapter is only the first step towards tackling structural factors constraining sustainable development.

(a)



(b)

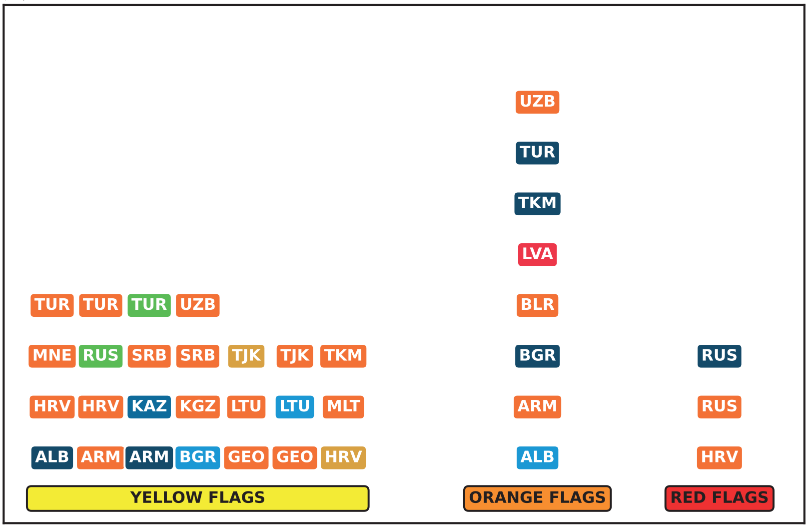


FIGURE 7.5 Bottleneck flags by country and SDG. (a) Africa, (b) Eastern Europe and Central Asia, (c) East and South Asia, (d) LAC, (e) MENA, and (f) West.

Notes: Each marker denotes an indicator with a bottleneck in a specific country (i.e., the name of the country is in the marker). We colour the markers according to the SDG to which the indicators belong. We identify these bottlenecks by simulating a relaxed budgetary frontier with $\gamma = 90\%$.

Sources: Authors' calculations with data from the 2021 Sustainable Development Report.

(c)

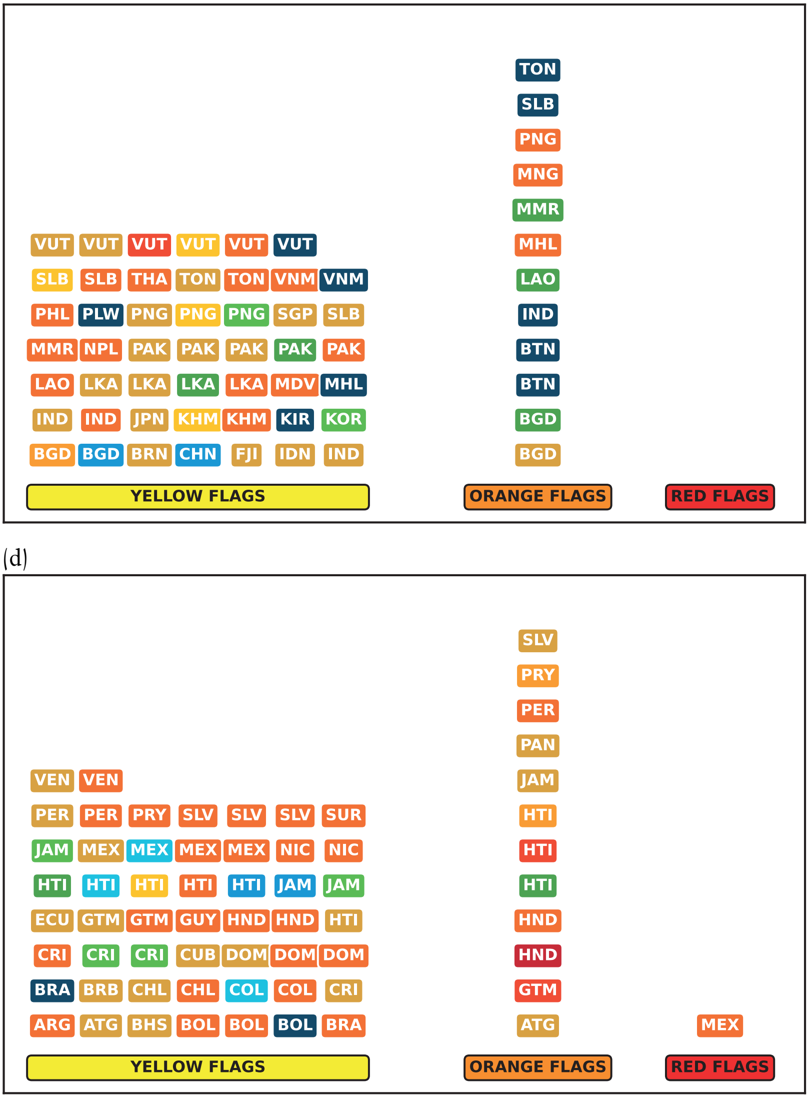


FIGURE 7.5 *(cont)*

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

FIGURE 7.5 *(cont)*

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

As we have previously explained, PPI cannot provide information about what is wrong in the government programmes associated with idiosyncratic bottlenecks since our model does not specify that level of detail. Therefore, when PPI identifies the potential bottlenecks, it is necessary to work with experts in those policy issues and, in this manner, broaden the scope of methods that could be employed to diagnose the effectiveness of programme-specific mechanisms. In other words, PPI can be easily coupled into a development diagnostics strategy and complement other methods by narrowing down the number of policy issues that need the government's attention.

7.4 SUMMARY AND CONCLUSIONS

The dire evolution observed in many SDG indicators around the world is not only a problem of funding availability and decreasing marginal returns, but it is also a consequence of structural bottlenecks. In this chapter, we have discussed and formalised one type of bottleneck relating to the structural constraints associated with a particular policy issue; we call this an idiosyncratic bottleneck. The presence of idiosyncratic bottlenecks implies that poorly performing indicators stagnate or even fall due to stringent long-term structural constraints. There are structural problems in a policy issue when the associated government programmes are ill-conceived or do not have a proper implementation. The first defect happens when the programme involves wrong incentives or collective action hazards. The second occurs because of logistic problems, inadequate human capital, absent infrastructure, or lack of maintenance in plants and equipment. Although we cannot identify the sources of these constraints with the model, we can use PPI to decipher the relative importance of long-term (structural) *versus* short-term (expenditure-related) effects in the dynamics of the indicators.

The evidence we collect from our counterfactual simulations shows that bottlenecks are widespread across SDGs and country groups. However, their relative importance varies among these groups. For instance, the existence of idiosyncratic bottlenecks in

https://doi.org/10.1017/9781009022910.011 Published online by Cambridge University Press

countries of East and South Asia is more common in comparison with other country groups. In contrast, these structural constraints are less prevalent in countries from MENA. Moreover, around the world, SDG 9 exhibits many bottlenecks, standing in sharp contrast to SDGs 4 and 8. In most cases, these long-term constraints only preclude the possibility of higher positive growth in the corresponding indicators. In a few more, they prevent moving from a negative to a positive tendency. In the remaining cases, the long-term constraints preclude the possibility of exhibiting a less severe negative trend. Consequently, we suggest a flagging system to differentiate among bottlenecks. On the one hand, researchers in sustainable development can use the proposed framework to understand the structure of the constraints dampening a country's potential. On the other hand, policymakers can use this information to prioritise public funding and identify government programmes that need a major revision.

Altogether, the results presented in this chapter show that the lack of public funding and structural constraints hamper the possibility of speedy advances across the SDGs that need to improve most urgently. Furthermore, these exercises illustrate, once more, the benefits of counterfactual simulations. By assuming a hypothetical steady positive growth due to an enlarged budget in the indicators' dynamics (i.e., the existence of a country operating under a budgetary frontier), it is possible to disentangle the consequences of long versus short-term effects in the indicators' evolution. Policymakers cannot always deal with the former effects because they require significant overhauls in the programs' incentives and operation procedures. Accordingly, these features usually require long and strenuous political discussions and statesmanship. Usually, such reformist endeavours entail institutional changes. Hence, in the next chapter, we investigate the problem of institutional change through the lens of public governance revision.