Subnational sustainable development: The role of vertical intergovernmental transfers in reaching multidimensional goals

Omar A. Guerrero [a,b,*], Gonzalo Castañeda [c], Georgina Trujillo [d], Lucy Hackett [e], Florian Chávez-Juárez [f]

[a] *Department of Economics, University College London, United Kingdom*
[b] *The Alan Turing Institute, United Kingdom*
[c] *Centro de Investigación y Docencia Económica, Mexico*
[d] *Centro de Estudios Demográficos, Urbanos y Ambientales, El Colegio de México, Mexico*
[e] *UC Berkeley, Department of Agricultural and Resource Economics, United States*
[f] *CORESO - Collaborative Research Solutions, Switzerland*

A B S T R A C T

From a public finance point of view, achieving sustainable development hinges on two critical factors: the subnational implementation of public policies and the efficient allocation of resources across regions through vertical intergovernmental transfers. We introduce a framework that links these two mechanisms for analyzing the impact of reallocating federal transfers in the presence of regional heterogeneity from development indicators, budget sizes, expenditure returns, and long-term structural factors. Our study focuses on the case of Mexico and its 32 states. Using an agent-based computational model, we estimate the development gaps that will remain by the year 2030, and characterize their sensitivity to changes in the states' budget sizes. Then, we estimate the optimal distribution of federal transfers to minimize these gaps. Crucially, these distributions depend on the specific development objectives set by the national government, and by various interdependencies between the heterogeneous qualities of the states. This work sheds new light on the complex problem of budgeting for the Sustainable Development Goals at the subnational level, and it is especially relevant for the study of fiscal decentralization from the expenditure point of view.

1. Introduction

The United Nations 2030 Agenda for Sustainable Development (aka Sustainable Development Goals or SDGs) was conceived through the lens of national-level data. However, around the world, many public policies associated with the SDGs are implemented at the subnational level, especially during the last two decades of the 20th century since many government programs have undergone a decentralization process [1,2]. For this reason, it is crucial to build databases with information at the regional/state/municipal/city level and to create analytical tools to support regional development decision-making and fiscal federalism.[1]

In addition to the pressing need of generating new data, sustainable development faces important operational challenges such as establishing budgetary priorities in a high-dimensional policy space (perhaps with hundreds of policy issues), accounting for the complex network of interactions between all policy issues, and dealing with inefficiencies in the use of public resources. All these features call for the creation of innovative analytical frameworks.[2]

In this paper, we apply a computational model to gain new insights into regional development and the decentralization of public expenditure. We assemble a dataset with 103 social, economic, and environmental indicators for 32 Mexican states.[3] The model facilitates

---

* Corresponding author. Department of Economics, University College London, United Kingdom.
E-mail address: oguerrero@turing.ac.uk (O.A. Guerrero).

[1] Formally speaking, fiscal federalism refers to the joint decentralization of taxation and government spending at the subnational level.
[2] The SDG literature recognizes that policy-making at the subnational level (states/municipalities/cities) is crucial if the 2030 Agenda is to succeed [3–7].
[3] Data availability for a large number of development indicators, years, and subnational entities is rare. Therefore, Mexico is one of the only countries through which we can systematically study the SDGs at a subnational level. Furthermore, its political economy and regional disparities have produced, in the last four decades, important fiscal imbalances that motivate the analysis of vertical intergovernmental transfers [8].

O.A. Guerrero et al.

estimating how feasible it is for the Mexican states to attain, by 2030, the goals established for the indicators. Our results indicate that the SDGs are not entirely viable, so we, subsequently, analyze whether convergence time can be reduced either by increasing the size of the state budgets (in relation to the empirical ones) or by modifying the nature of federal transfers (a critical component of the states' income). We find that, on the one hand, the sensitivity of the indicators' performance to budgetary increments is heterogeneous across policy issues and states. On the other hand, our simulations indicate that significant reductions in convergence times can be achieved when an 'optimal fiscal transfer' is used to allocate the federal transfers across subnational central governments (SCGs).

The latter result is particularly important because fiscal federalism has different flavors in different countries. From a theoretical perspective, it is argued that allocating spending powers to SCGs improves the provision of public goods and services, and helps to reduce personal and regional income disparities.[4] However, to avoid double taxation and to attain economies of scale in tax collection, national governments are reluctant to grant full taxation autonomy to SCGs. Because this perspective creates vertical fiscal imbalances, some form of decentralization of revenues should be promoted. Among other things, it is argued that a certain leeway in tax autonomy tends to encourage local fiscal capacity and fosters a competitive environment that is conducive to the improvement of social and public infrastructure.

Different forms of fiscal federalism demand the use of flexible analytic tools to facilitate their evaluation.[5] In this paper, we employ an agent-based model to study how fiscal revenues can be shared across subnational entities to reduce the states' average development gaps expected by 2030. This systemic approach is very helpful since traditional econometric strategies only consider one measure of performance at a time, for example, poverty alleviation, inequality, regional convergence, or economic growth. Furthermore, existing studies typically look at highly specific forms of decentralization, such as transfers in health, education, and social infrastructure [12–16]. Needless to say, understanding how governments allocate their resources to reach development goals comes with serious endogeneity issues because the outcome variables–the indicators–often inform budget-allocation decisions. As we show in this paper, agent-based models allow overcoming some of these problems and enable new ways to analyze decentralization in public spending.

For a better understanding of the model's implications, the reader should be aware that our inference of policy priorities is derived only from budgetary considerations. That is, when performing counterfactual experiments, the simulations assume that the same policies implemented in the past remain in place. Hence SDG feasibility is defined exclusively in terms of the size and allocation of resources. The model is not designed to produce *ex ante* evaluations of micro-policies in which incentives, externalities, and operational attributes are modified to close development gaps. Furthermore, no criticism is made concerning the practicality of some of the SDGs. In our model, the goals established by governments are exogenously given, hence our objective is to analyze if such goals can be attained with the existing policies, and how public resources should be allocated across government programs and SCGs.

The remaining of this paper has five more sections, structured in the following way. Section 2 reviews the main features of the Mexican fiscal federalism from the expenditure side. Section 3 provides an overview of the model and of the relevant literature. Section 4 describes the nature of the Mexican database and how the goals were established for each of the SDG indicators. Section 5 shows summary statistics and visualizations of

our results. The results describe the feasibility of the goals and how much savings in convergence time can be obtained by increasing the budget or by optimizing the distribution of federal transfers. Section 6 presents some discussions and conclusions.

2. Fiscal decentralization in Mexico

It is convenient to briefly discuss some characteristics of the Mexican fiscal federalism from the point of view of public expenditure. Mexico has a deep vertical fiscal imbalance, in so far as its political economy has produced a weak fiscal capacity at the state and municipal levels and unequal decentralization ratios. For example, subnational tax revenue was close to 5% of the total government tax revenue in 2017, while subnational spending was about 40% of total government spending. This imbalance has created the need to share considerable sums of fiscal revenues across 32 states with widely different political powers and economic structures.

Through the Fiscal Coordination Act of 1978, substantially reformed in 1998 and 2007, the Mexican national government distributes fiscal resources through two main channels: 'participations' (unconditional transfers known as *participaciones*)[6] that can be used according to the states' objectives, and 'contributions' (conditional transfers known as *aportaciones*).[7] Contributions are tied legally–but not always in practice–to broad activities such as health, education, and social infrastructure. The overall size of the participations corresponds to 21% of the Shareable Tax Revenues of the Federation (STR), which is composed of income tax, value-added tax, special taxes on goods and services, and taxes on oil and mining extraction.

From the perspective of the national budget, the participations are consolidated in the so-called *Ramo 28*, a tranche of shared revenues whose main objective, according to official documents, is to compensate state governments for federal taxes collected within their territories. These revenues are transferred to the governments of the 32 states using a formula that, according to different analysts, relies heavily on population size [17]. The contributions, on the other hand, appear in the national budget as the so-called *Ramo 33*, a tranche of shared revenues comprising[8] different funds, each one aimed at equalizing regional disparities in specific dimensions of development (e.g., health and education). The procedure to allocate these transfers varies from one fund to the other. Approximately 29% of the STR is directly transferred to states and municipalities, while the residual is managed by the federal government.[9]

Fig. 1 presents the state budgets approved for 2019, disaggregated into locally collected revenues and the federal transfers. Federal transfers, in turn, are sub-divided into participations and contributions. Notice that, in all states, federal transfers are much larger than local revenues. Participations tend to be relatively large in states with a high GDP such as Mexico City (CMX), Nuevo León (NLE), and the State of Mexico (MEX). Contributions, in contrast, dominate in states with low GDP such as Oaxaca (OAX), Chiapas (CHP), and Guerrero (GRO). Only the State of Mexico and in Mexico City have the capacity to procure a

---

[4] See Oates [9] for arguments in favor of fiscal decentralization, and Prud'-homme [10]; Tanzi [11] for arguments against it.

[5] See a meta-study for OECD countries by Forman et al. [1]; and an analysis for Latin America by [37] to learn more about various types of fiscal federalism.

[6] While the literal translation of *participaciones* is 'shares', we use the adapted term 'participations' to avoid confusion with the shares of a budget and other related concepts. Moreover, in the English-speaking literature of Mexican fiscal federalism, participations are also referred to as 'shared revenues', but this alternative term does not differentiate between conditional and unconditional transfers.

[7] Participations and contributions correspond, approximately, to 80% of the revenues shared by the federal government.

[8] For more details on the Mexican system of fiscal coordination, the reader can consult the following studies: Diaz-Cayeros [8]; Sánchez and Ballinez [18]; Chiguil [19]; Hernández and Rabling [20,36].

[9] For more details on the Mexican system of fiscal coordination, the reader can consult the following studies: Diaz-Cayeros [8], Sanchez and Ballinez [18], Chiguil [19], Hernandez and Rabling [20], Giugale and Webb [36].
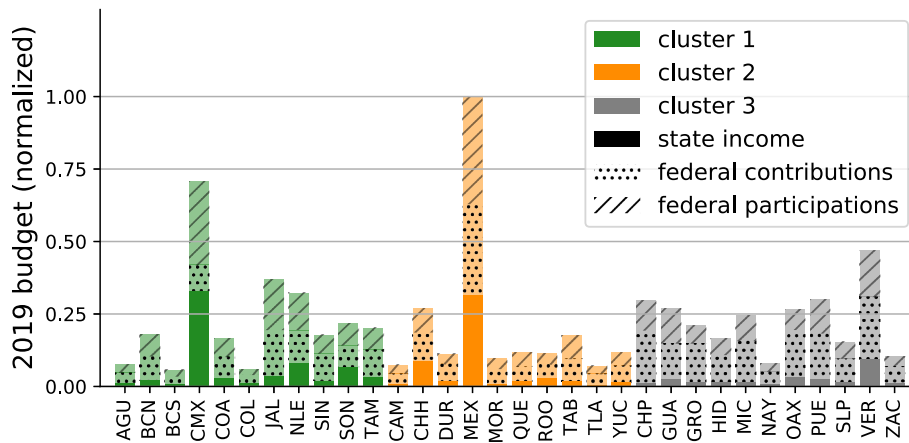
Fig. 1. State budgets in 2019. *Note*: The total budgets have been disaggregated into state income–solid bars–and federal transfers–translucent bars. The latter, in turn, are decomposed in contributions)–doted bars–and participations–stripped bars. Details on the state names and clusters are provided in section 4. *Source*: INEGI (inegi. org.mx/programas/finanzas).

significant local tax collection. On the contrary, the states of Chiapas (CHP), Guerrero (GRO), Michoacán (MIC), Morelos (MOR), Puebla (PUE), and Tlaxcala (TLA) have extremely weak bases for the generation of local revenue. Altogether, the data reveal unevenness in intergovernmental transfers and, thus, their distribution affects the states' possibilities to reach the SDGs.[10]

3. A brief explanation of the model

We develop an agent-based model of the policymaking process that allows linking public spending to policy outcomes (measured through development indicators) under a causal framework. The model is based on Guerrero and Castañeda [21]; who study the feasibility of the SDGs worldwide. Guerrero and Castañeda [22,23] provide a thorough motivation of a closely related model, and an application to the national case of Mexico. Earlier works using similar models can be traced back to Castañeda et al. [24]; who focus on socioeconomic indicators. Other precedents with applications to different problems can be found in Castañeda and Guerrero [25], who study the resilience of development when facing sudden disruptions in specific activities; Castañeda and Guerrero [26], who use a similar agent-based model to perform *ex ante* evaluations of specific budgetary allocations; Guerrero and Castañeda [22,23], who quantify policy coherence; and Guerrero and Castañeda [27], that evaluates the role of public governance mechanisms in the fight against corruption.[11]

We provide the full details of the model in Appendix B, and highlight the most relevant features for this application here. For an outline of these components, we provide a diagrammatic sketch that explains the overall logic of the model. It is important to mention that this is different from more popular data-fitting exercises such as estimating regression coefficients or machine learning models (recall that these approaches cannot deal with the aforementioned endogeneity issues, and neither with the complex interdependencies between indicators). Instead, an agent-based approach seeks to generate highly-granular data on the

individual behavior of the agents. These data are then aggregated into variables that correspond to real-world ones. Thus, the calibration procedure seeks to find a set of deep parameters that allow a minimal error between certain empirical features and their simulated counterparts. In this paper, those features are the values that each indicator achieves at the end of the sampling period, and the average growth rate of the indicators. Thus, once calibrated, the model offers a bottom-up causal account of the observed dynamics of the 103 indicators across the 32 states.

Now, let us proceed by explaining the sketch presented in Fig. 2. The model simulates the behavior of a government (or central authority) that assigns public funding to different programs (or policy issues) intending to reach a set of goals. Such goals are defined exogenously in the model and, ideally, they describe the government's vision on what development path it aspires to take.

The upper left circle of Fig. 2 represents a government whose budgetary decisions are incentivized by the desire to reach a pre-established set of goals (lower left circle). Each period, this central authority allocates resources to functionaries (upper-middle circle) in charge of implementing different government programs (one program/ agent per indicator). The latter agents have a mandate to use the assigned funds for improving the performance of the associated indicators (lower right circle). However, once the allocated funds are received by the public servants, there can be inefficiencies of different sorts (e.g., embezzlement of resources for personal use, excessive bureaucracy, mismanagement of goals, or ill-conceived public tenders). Exercising these funds across the ecosystem of government programs creates spillover effects due to network interdependencies, which are exogenous and are an input of the model (upper right circle). The direct and indirect effects of these resources impinge upon the whole spectrum of SDG indicators (lower right circle), and this could happen through negative spillovers (trade-offs), or positive ones (synergies).

The waste of resources produced by the misalignment of incentives between the central authority and its functionaries can be ameliorated through regulatory and judicial schemes, which are exogenous parameters, imputed from empirical data. The model recognizes that, in reality, these schemes tend to be imperfect and may depend on the quality of the available tools of public procurement (e.g., rule of law and the monitoring of corruption). Accordingly, public servants learn, through time, how much inefficiency is tolerated in the political system (upper dark blue small horizontal arrow). While the initial allocation of resources is established as an attempt to reduce development gaps across the different indicators (lower long line), as time goes by, the government reacts by adapting such allocation according to the policies' perceived inefficiencies by rewarding the most efficient policymakers

---

[10] These fiscal transfers are not the only channels that the federal government uses to spend in states and municipalities. There are also matching grants for projects in public infrastructures and social assistance programs, such as the well-known programs of *Procampo* and *Oportunidades*. The cash-transfers of these two programs are assigned directly to individuals by the federal administration and, thus, their resources are not under the control of subnational governments.

[11] A complementary publication to this literature can be found in Ospina-Forero et al. [28]; where an extensive analysis of network-estimation methods for SDGs is provided.
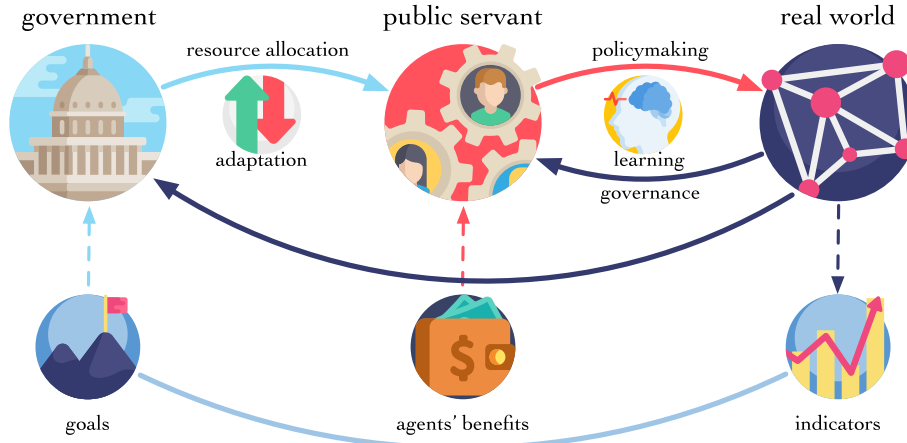
Fig. 2. Model sketch. *Source*: [22,23].

(middle dark blue long horizontal arrow).

For further clarity, it is important to explain two mathematical formulations that are part of the macro-components of the model. First, we have a system of evolution equations that characterize indicator dynamics. The difference equation of indicator i is specified as

$$I_{i,t+1} = I_{i,t} + \alpha_i \xi(\gamma_{i,t}), \qquad (1)$$

where $I_{i,t+1}$ is the level of indicator i at time $t + 1$; α_i is a growth factor that incorporates long-term structural features not explained explicitly by the model; $\xi(\cdot)$ is a binary variable that takes the value of 1 when a positive growth event is realized, and a value of 0 otherwise; and $\gamma_{i,t}$ is a probability function to be explained below. This formulation is composed of an inertial term and a growth term. The latter explains the transformative capacity of the indicator. It depends on the interaction between long-term structural factors (α) and short-term outcomes from implementing already-existing government programs, so it may be sensitive to changes in the financing of those programs.

Second, we define a set of probability functions that describe the likelihood of a positive growth event. These equations are specified as

$$\gamma_{i,t} = \beta \frac{C_{i,t} + \frac{1}{n}\sum_j C_{j,t}}{1 + e^{-S_{i,t}}}, \qquad (2)$$

where β is a normalizing parameter capturing, among other things, the productivity of the resources invested in government programs associated with indicator i[12]; $C_{i,t}$ is the amount of resources effectively used (after inefficiencies) by the ith public servant in period t; and n is the number of instrumental indicators (see section B for the definition of instrumental indicators). The term $S_{i,t} = \sum_j \mathbf{1}_{j,t}\mathbb{A}_{j,i}$ represents the net incoming spillovers received by indicator i in period t. $\mathbf{1}_{j,t}$ is an indicator function that returns 1 if indicator j grew in the previous period and 0 otherwise, and the adjacency matrix $\mathbb{A}_{j,i}$ corresponds to the network of

interlinkages, with each entry representing a conditional dependence calculated according to the Bayesian method established by Aragam et al. [29].[13] This formulation indicates that the probability of a successful growth event relates positively with the size of resources effectively used in a public policy and with the magnitude of positive spillovers coming from other indicators. Negative spillovers, in contrast, reduce the likelihood of growth in an indicator.

Appendix B provides information regarding further microeconomic details about the incentives of the agents, their learning rules, the monitoring of inefficiencies, their penalization, and the government's heuristic to adapt its allocations.[14] None of these detailed components contains free parameters other than the ones already presented in equations (1) and (2). Thus, the parameters that need calibration are α_1, ..., α_N, β, where N is the total number of indicators. The calibration is performed for each state independently by using an algorithm developed by Guerrero and Castañeda [21]. Broadly speaking, the algorithm seeks to find a vector $\alpha_1, ..., \alpha_N$ that minimizes the difference between the final value of each empirical indicator, and the ones expected from a set of Monte Carlo simulations. At the same time, it finds a parameter β that minimizes the difference between the mean probability of success $\gamma_{i,t}$ endogenously generated by the model (averaged across indicators, periods, and simulations) and the rate of positive growth periods across all

---

[12] It is important to highlight that β helps to correct for the influence of the population size in the budget of a state. On the one hand, population effects are removed from the indicators by taking per-capita rates. One the other, budgets could also be normalized per capita, but it would be inconvenient for the type of analysis presented in section 5.3, in which resources are reallocated across states. Instead, the budgets are left in their total values, and the calibration seeks to find a β that normalizes these data so it can work well as part of equation (2) in terms of generating probabilities. In other words, if two states have similar indicator performances, but one has a larger budget due to its population, then the β of the more populous state has to be smaller to generate a probability $\gamma \in [0, 1]$. Thus, β reflects that it is not the same to spend one peso in a more populous state than in an uncrowded one. This can be confirmed in Appendix D, where all the parameters β are plotted against population data.

[13] Consult Appendix C for more details on the interpretation, estimation, and implications of SDG networks.

[14] While GDP-focused models make an effort in modeling consumer behavior, they need to sacrifice realism in terms of the number of development dimensions that they can incorporate, as fully specifying every micro-level mechanism of the SDGs is unfeasible. Our model locates on the other side of this trade-off. It focuses on the micro-mechanisms of the policymaking process and provides a more stylized–macro–view of the SDGs. This is mainly motivated by our need to establish a causal link between government expenditure and development while considering all the dimensions of the latter contained in the SDGs. But even with this stylized view, when the model is calibrated, we can capture the aggregate outcome of micro-level decisions through the network's links and weights as well as the free parameters of the SDG indicators' evolution. For example, a negative link (trade-off) between an indicator in an economic SDG and an environmental one reflects a negative externality produced by firms' erroneous incentives and the available polluting technology. On the other hand, a small α reflects structural long-term considerations that hamper the potential of public funds on attaining development goals when government programs are based on ill-conceived consumers' and firms' incentives.

the empirical indicators of a given state.[15]

4. The data

In collaboration with the National Laboratory for Public Policy (LNPP),[16] we collected data from a variety of sources, including government agencies and NGOs within Mexico and international organizations such as the World Bank and United Nations. We build a balanced panel of 103 indicators for the 32 Mexican states from 2005 to 2018. We provide the full definition of all the variables, their source, and their SDGs in Appendix A.1. For budgetary data, we use information on revenue and expenditures from INEGI that is disaggregated by source, and that identifies state-collected revenues from federal transfers. See Appendix A.2 for more details on these data and their normalization.

Finally, we perform a manual classification of the indicators into 'instrumental' and 'collateral'. This is a classification created in earlier applications of related models, and allows identifying those indicators that can be directly impacted by existing government programs (instrumental), and those that are too aggregate to receive such directed impacts (collateral) [22,23]. Policy priorities can only be defined over instrumental indicators, so collateral ones improve through spillover effects and inertial factors. Once we have taken the previous steps, we consolidate the indicators with the budgetary data in a panel.

4.1. Quantification of development goals

To define the numeric goals of each indicator, we apply a combination of methods. Firstly, whenever possible, we base the goals on those defined by OECD [30]; which uses regional-level data (in the case of Mexico, state-level). These goals are adopted whenever the indicator has a close counterpart in the OECD data.[17] We use this method for 17 indicators. Secondly, when there is no clear equivalent between an indicator and the OECD data, we follow the protocol established by Lafortune et al. [31]; who define 5 methods for setting goals. This procedure represents a sequential decision tree. Thus, an indicator's goal is determined according to the first relevant criterion. However, we consider only four of these criteria[18] (see Table A2 for the full list of indicators and the method used to assign a goal to each one):

● **Criterion 1** sets the goal to an SDG target when the indicator describes a variable whose definition explicitly implies the existence of a quantitative threshold to be reached. For example, one of the SDG targets is to eradicate extreme poverty; thus the value assigned for the goal of indicator 'Percent of the population in extreme poverty' is set to 0. We also include some indicators that are not part of the SDG

ensemble, but that have explicit targets in Mexico's National Development Plan (NDP). For example, for the indicator 'Fertility rate of women aged 15–19', we set the goals to half the most recent level, as stated target of the NDP. We use this criterion to set goals related to gender or ethnic inequality by using the value that represents full equality. We apply this criterion to 9 indicators.

● **Criterion 2** sets the goal to an implied value based on the principle of 'Leave no one behind'. For example, we set a goal of 1 for the indicator 'Proportion of 1-year-old infants with a full basic vaccination record'. We apply this criterion to 11 indicators.

● **Criterion 3** sets the goal to an optimum value defined by scientific evidence. For example, we set the optimum number of policemen per 100,000 inhabitants to the worldwide median. This criterion is applied in 1 indicator.

● **Criterion 5** sets the goal to the average value of the top 5 states according to their best year of performance in the data. For a state that is above this average, we set the goal to its most recent value, assuming that its performance in such SDG indicator has already been successful and no gap needs to be closed. We apply this criterion in the majority of indicators: 65.

4.2. Development clusters

Our simulations are specific to each state. However, for clarity of exposition, we present some summary visualizations where we may aggregate the results into clusters of states or into SDGs. Importantly, the reader should not interpret such visualizations as proper estimates because all the results have been obtained at the state level.[19] Instead, they should be considered as stylized descriptions of outputs that are difficult to communicate along 103 indicators and 32 geographical entities. In terms of state clusters, we define three that represent groups of states with different levels of development. Using the Human Development Index (HDI), we construct these clusters according to those states below the 33.33 percentile, between the 33.33 and the 66.66 ones, and above the 66.6 percentile.[20] Fig. 3 shows the geographical distribution of the three clusters across the Mexican territory. This clustering provides some context about the heterogeneous levels of development across states, which should help the reader interpreting our results. Notice that, except for Mexico City, all states in cluster 1–colored in green–are located in the northern and pacific regions. In contrast, the states with the lowest human development–colored in grey–are positioned in the central and southwest regions. This indicates that neighboring states tend to share many affinities, social and otherwise (validating out data-imputation procedure).

To explain Fig. 4, let us define a retrospective gap as the total progress made by an indicator in the historical data, measured as the difference between the final value (in 2018) and the initial one (in 2005). A prospective gap, in contrast, is the difference between the development goal established for an indicator and its value in 2018, i.e. it is the progress that would be needed in 12 years to reach an SDG. The figure shows, in three panels, retrospective gaps in solid bars (in percentage) and prospective gaps in translucent bars (also in percent changes), both averaged across states within the corresponding cluster. The dots on the top of the translucent bars correspond to the development goals. Thus, a white dot indicates that, on average, the prospective

---

[15] The reader may ask why should one use this specific calibration algorithm, while there exist many methods for this purpose. Due to the interdependencies between indicators through the network and the endogenous responses of the agents, the fitness landscape of this optimization problem is dynamic and rough. In fact, conventional numerical methods for convex optimization fail. Heuristic methods such as differential evolution, and even Bayesian approaches such as the Tree-structured Parzen estimator also fail. The method developed in Guerrero and Castañeda [21] employs a multi-output gradient descent algorithm. The calibration is simultaneous for all parameters, so the method is efficient and minimizes indicator-specific errors; and Guerrero and Castañeda [21] show that it yields high fitness scores.

[16] The LNPP (https://www.lnpp.mx) is a think-tank located in Mexico City that, among other tasks, maintains a panel of hundreds of indicators at the state level.

[17] Note that the OECD dataset has very short time series, often with only one or two observations for the indicator of a given state. Therefore, we do not include any of this information in our database.

[18] We omit criterion 4, which takes into account values of countries that have already surpassed the official SDG target, as international values may not be comparable to state indicators.

[19] Because we calibrate the model for each state individually, and perform our simulations also for each state independently, concerns about potential biases due to the grouping of states or indicators are invalid.

[20] The HDI was designed to measure multidimensional development (health, education, and standard of living) and to compare well-being between countries. To display differences within countries, a Subnational Human Development Database project was conducted by Smits and Permanyer [32]. This Subnational Human Development Index has been constructed for 161 countries and includes Mexico for the period 1990–2018.

HID: Hidalgo
JAL: Jalisco
MEX: State of Mexico
MIC: Michoacan
MOR: Morelos
NAY: Nayarit
NLE: Nuevo Leon
OAX: Oaxaca
PUE: Puebla
QUE: Queretaro
ROO: Quintana Roo
SLP: San Luis Potosi
SIN: Sinaloa
SON: Sonora
TAB: Tabasco
TAM: Tamaulipas
TLA: Tlaxcala
VER: Veracruz
YUC: Yucatan
ZAC: Zacatecas

AGU: Aguascalientes
BCN: Baja California
BCS: Baja California Sur
CAM: Campeche
COA: Coahuila
COL: Colima
CHP: Chiapas
CHH: Chihuahua
CMX: Mexico City
DUR: Durango
GUA: Guanajuato
GRO: Guerrero

Fig. 3. State clusters according to the Human Development Index. *Note:*Green states: upper tier of the HDI (cluster 1). Orange states: middle tier of the HDI (cluster 2). Grey states: bottom tier of the HDI (cluster 3). *Source*: Smits and Permanyer [32] and authors' own calculations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

gap is smaller than the retrospective one, i.e. that the indicator had a better historical performance than the one needed to reach its goal by 2030.

The solid bars show that most of the indicators experienced positive growth during the sampling period. However, there are some indicators in SDG 16 with important drops, especially in cluster 3. Likewise, the slight fall observed in the ecological indicators of SDG 15 is more notorious in clusters 1 and 2. In these panels, we can see that the gaps are heterogeneous across indicators, irrespective of the development cluster. The patterns of the translucent bars for the three clusters are rather similar; nonetheless, there are some differences between panels. For instance, there are 8 indicators with goals well above 40% in cluster 3 (where the least advanced states are grouped) but only 4 and 3 indicators in clusters 2 and 1, respectively. Should the reader wish to make more nuanced comparisons between the level of development of the different states and indicators, see Appendix A.

5. Analysis of the simulations

Once we calibrate the model parameters, we produce three types of analyses. Firstly, we are interested in studying temporal feasibility; that is, the progress that is likely to take place with regards to the SDGs by 2030. Secondly, we analyze the sensitivity of the indicators' performance when the generalized budget growth or shrinkages. The *ex ante* evaluation of these scenarios is relevant under different economic settings. For instance, an economy can undertake a fiscal reform to enhance tax revenues; there can be windfall gains when the exports of natural resources experience a temporal surge; a negative shock–such as a pandemic–can deteriorate the health of public finances; and demographic pressures can enlarge liabilities associated to the payment of government pensions. Thirdly, we explore the potential outcomes of intergovernmental fiscal transfers as a function of the goals. While fiscal federalism is a reality in many countries, and Mexico is not the

exception, the literature does not facilitate a clear understanding of the multidimensional consequences of decentralizing government expenditure. Thus, these last exercises seek to shed new light on the links between multidimensional development goals, indicator performance, and decentralized public spending.

5.1. SDG gaps by 2030

Our first set of results consists of the estimated gaps between the level of the SDG indicators in 2030 and the proposed goals for that year. To obtain these estimations, we perform 1000 Monte Carlo simulations for each state. Each simulation runs forward in time, starting with the values of the indicators in 2018 and stopping when the simulation reaches 2030. The budget size of each state corresponds to the one approved for the fiscal year of 2019, projected over 12 years. An SDG gap consists of the difference between the average (across simulations) final simulated value of each indicator (of a given state) and its corresponding SDG value, divided by the goal and presented in percentages. More specifically, we compute

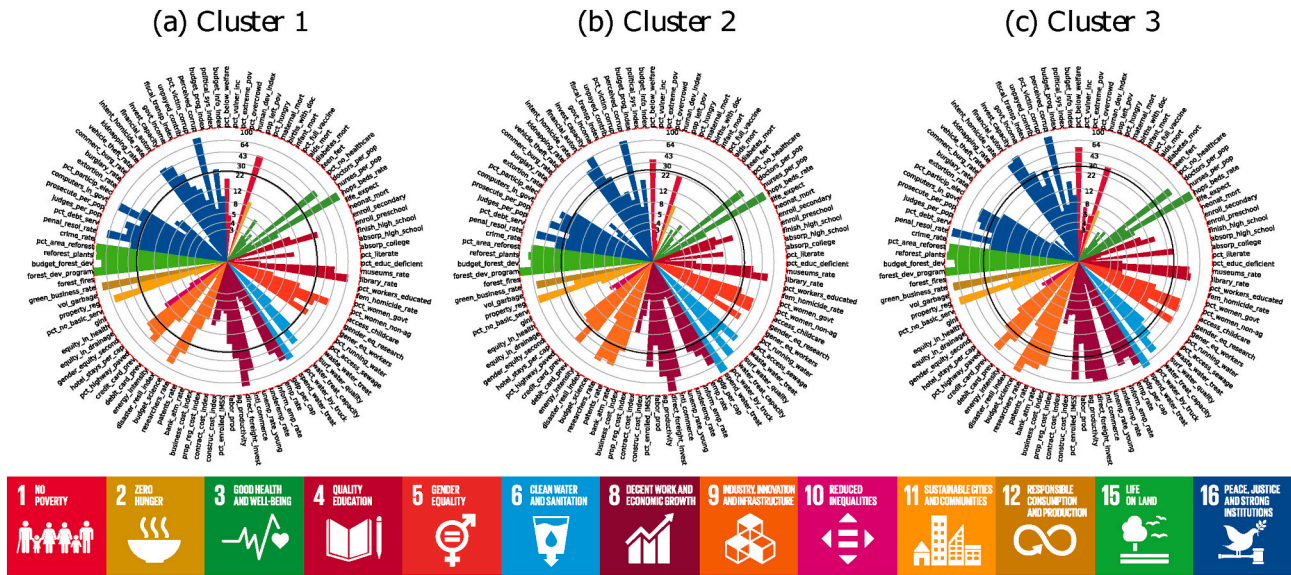$$SDG_{gap}^i = 100 \times \frac{\max(0, SDG_i - I_{i,T})}{SDG_i}, \tag{3}$$

Fig. 4. Retrospective and prospective gaps. *Note*: The solid bars denote the net change between the final and initial values of the indicators as a percentage of the initial value, averaged across states in the same cluster. The translucent bars indicate the difference between the value of the indicator in 2018 and the SDG to be achieved as a percentage of the 2018 value, averaged across states in the same cluster. The dots indicate the SDG of the indicator. If a dot does not appear in the plot, it is because the visualization cropped it out.
Source: Authors' own calculations.

where i denotes the indicator and T is the number of years.[21] Thus, the

SDG gap for $T = 12$ can be interpreted as the gap–in percentage terms–that has not yet been closed in 2030 for indicator i in a given state. Appendix D provides plots with the parameters obtained through the calibration procedure.

Fig. 5 presents the SDG gaps aggregated into clusters. The first thing to notice is the fact that, independently of the cluster, few indicators–averaged across simulations and states–can reach their goals by 2030 (they exhibit gaps of less than 5%). The estimated average gap across indicators is 22% for the states in cluster 1, and nearly 27% for states in cluster 3. This visualization illustrates that many indicators in SDG 16, associated with 'Peace, Justice and Strong Institutions', make only modest improvements; and the same can be said about ecological indicators in SDG 15. While there is a similar pattern across the three

---

[21] Note that the number of simulation periods is different from the number of years. When calibrated, the model must run for at least the number of years in the data. Guerrero and Castañeda [21] show that 50 or more simulation periods yield robust results (because the budget parameter β adjusts for the volatility of the indicators). Thus, once a specific number of simulation periods is chosen, a proportional equivalence between this amount and the number of years in the dataset is established. Therefore, when performing prospective simulations, the number of simulation periods to be used must hold according to the established equivalence of the calibration. In this application, we establish 50 simulation periods for 14 years of data. Thus, the number of simulation periods required to run 12 years forward in time is approximately 43.

ARTICLE IN PRESS

O.A. Guerrero et al.
Socio-Economic Planning Sciences xxx (xxxx) xxx

Fig. 5. Average gaps by 2030. *Note*: The height of each bar denotes the average gap between the SDG and the indicator level predicted by 2030 and computed across states in the cluster. The solid black ring corresponds to the average gap across states (in the cluster) and indicators. A dashed red ring refers to the largest average gap (between indicators in the cluster).
Source: Authors' own calculations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

clusters, there are salient differences as well. For instance, the SDG gap for 'Workers enrolled in IMSS as a percent of the total population'[22] is below the cluster average in the most developed states, but not in clusters 2 and 3. The same happens for 'Percent of young people aged 19–29 with income below the welfare line' and for 'Rate of informal employment'.

Two implications are derived from these simulation results. Firstly, SDG gaps will hardly be closed by 2030, even for the richest states belonging to Cluster 1, assuming that their annual budgets remain at their historical levels in real terms. Secondly, we cannot expect regional convergence shortly if disruptive policies are not implemented, since SDG gaps tend to close more sluggishly in states classified in Cluster 3. Likewise, the problem of lack of convergence varies in terms of the indicators analyzed, being especially important for social security coverage, poverty, and labor informality. Appendix E provides a more nuanced visualization of the development gaps by state and indicator.

5.2. Sensitivity to changes in the budget size

While state-level heterogeneity may give rise to different SDG gaps, it is unclear how these gaps change as a result of the budget size. That is, if the expenditure level of a state were to increase, would this necessarily drive an improvement in all its indicators? Would such a relationship be linear? Would all indicators improve at the same rate? Would the indicators worsen in a symmetric pattern if the budget were reduced? All these questions are critical for understanding the link between public financing and regional development. In this section, we answer them through simulation experiments.

Let us begin by comparing the expected (average) level that each indicator achieves by 2030 (obtained through the simulations that produce Fig. 5) and those with a 1% increase in the budget size. Then, taking into account the possibility that responses to budgetary changes may be non-linear, we repeat the previous procedure, but considering the effect of a budgetary increment that goes from 1% to 2%. Afterward, we perform a series of similar exercises (in increments of 1%) until we

reach a 50% increment. Accordingly, the average change across all these marginal increments is the estimated change in the indicator when the budget increases by 1%. Similarly, we estimate the indicators' decline resulting from 1% reductions of the budget (reaching a 50% reduction). We perform these two calculations for each state and present the disaggregated results for each indicator in Fig. 6.

The first thing to notice when comparing panels (a) and (b) is that, on average, decrements in the budget exert a stronger impact than increments do. This asymmetry is a reflection of equations (2) and (1), which suggest that a better financial situation is a necessary condition for development (otherwise $\gamma = 0$) but not a sufficient one (since growth is always limited by α). Moreover, in clusters 1 and 2, but not in cluster 3, several indicators show extreme increments/decrements (red dots in the diagrams). This result indicates that the impact of budget size is less heterogeneous for the least developed states of cluster 3. Likewise, from the bottom panel, it is clear that the ecological indicators of SDG 15 (light-green dots) deteriorate when public funds shrink. Finally, the dots in each column suggest that outcomes differ between states, which highlights the importance of context specificity.

Analyzing convergence times makes evident the non-linear response of the indicators to budgetary changes. That is, for each indicator that arrived at its SDG by 2030 or before, we record the number of years that it takes to reach its goal.[23] Then, we compute the convergence time when applying a series of 1% increments and decrements to the state budgets previously considered. For each calculated time, we compute the difference with respect to the benchmark simulation (the one with the historical budget of each state). Thus, the differences that result from an enlarged budget represent years saved, while the ones originating from a budget shrinkage indicate delays. By aggregating these saving and delays across states in the same cluster and across indicators in the same SDG, we construct convergence-sensitivity curves at the level of SDGs and clusters, which we present in Fig. 7.

Figs. 6 and 7 share four features: asymmetric impacts for budgetary increments and decrements, heterogeneous sensitivity across clusters, negative impacts on the ecology-related SDGs caused by budgetary
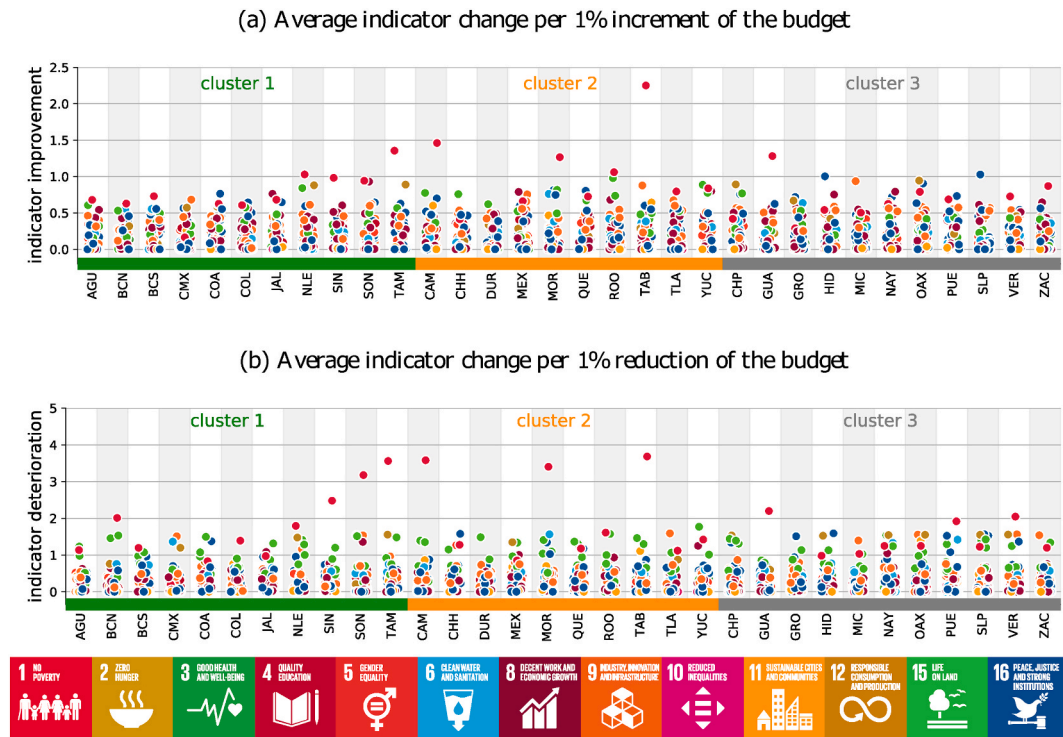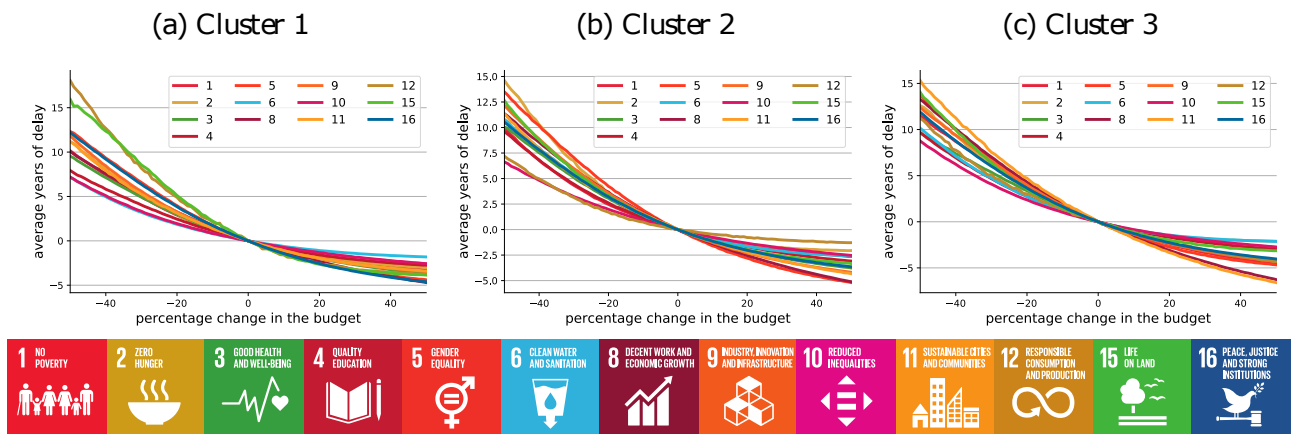

Fig. 6. Sensitivity of indicator levels to changes in the size of the budget. *Note*: Each dot corresponds to the average response of a single indicator to a 1% change in the budget size of the corresponding state. This average value is computed across 1000 Monte Carlo simulations for each percent change in the [1%, 50%] range. Panel (a) presents the results for budgetary increments, while panel (b) shows the result for decrements. *Source*: Authors' own calculations.

Fig. 7. Sensitivity of convergence times to changes in the size of the budget. *Note*: Each curve consists of the aggregated responses in convergence times of indicators in the same SDG and cluster to 1% budgetary changes. These curves can only be constructed for indicators that reached their goals in or before 2030. Thus, the specific quantities may result from a selection bias in the indicators. This is not an issue in Fig. 6. Here, the reader should rather focus on the qualitative nature of the results. *Source*: Authors' own calculations.

reductions (especially in cluster 1), and discernible region-specific patterns. Notice the diminishing sensitivity to budget increments visualized by the asymptotic blue lines–from SDG 6–in clusters 1 and 3, and the golden line–from SDG 12–in cluster 2. They reinforce the idea that enlarging the budget is not always enough for accelerating development. The convexity of these sensitivity curves and the fact that it is common to find crossings–especially on the budget reduction side–indicate non-linearities. While these line shapes are largely driven by equation (2), the non-monotonicity and crossings can be attributed to the dynamics associated with the learning and adaptation of the agents.

The most important implication derived from the previous set of simulation results is that indicators do not respond significantly to sizable changes in the overall budget in any of the Mexican states,

independently of how rich or poor they are. These results illustrate that the role of public spending to foster development is limited, even if enough public funds were available. Hence, disruptive micro-level policy interventions are needed to replace many of the existing government programs. The diminishing returns of policies based entirely on the size and allocation of the budget will hardly allow the least developed states in Mexico to move towards a stronger path.

5.3. Analysis of intergovernmental transfers

A critical issue that countries face when undertaking reforms towards fiscal federalism is determining how public expenditure should be decentralized. In particular, the allocation of intergovernmental fiscal

transfers is a process that deserves further study. We propose a simulation method to produce *ex ante* evaluations of allocations resulting from emphasizing broad categories of development goals. The allocation of transfers can be an empirically observed distribution or a hypothetical one, based on discretionary decisions or pre-determined rules. While our method is agnostic about the type of objective functions to be considered, here we propose an application that focuses on the SDGs. This approach is flexible enough to consider multiple dimensions of development simultaneously.

Equation (4) presents a comprehensive objective function–or fitness landscape–to be minimized by a heuristic optimization algorithm known as differential evolution. This method is particularly well-suited to work under rough high-dimensional fitness landscapes. The objective function consists of a weighted average (across states) of all SDG gaps. Its summands correspond to different indicators and are calculated according to equation (3). The weight that ponders the contribution of a state is determined according to population size. The absence of such weights would imply that extremely small states would be assigned unnecessary fiscal transfers. However, population-based weights do not imply that the number of inhabitants is the determining factor when assigning intergovernmental transfers, as is shown by

$$F = \sum_{j=1}^{32} \omega_j \sum_{i=1}^{N} SDG_{gap}^{i,j}, \tag{4}$$

where ω_j is the population weight for state j.

A refinement of this objective function consists of isolating only a subset of indicators that are considered relevant as criteria to optimize the distribution of federal transfers. In the context of the SDGs, such subsets could be naturally constructed by focusing the optimization algorithm on those indicators that belong to a given SDG. Thus, we define the objective function for an SDG-specific optimization as

$$F_k = \sum_{j=1}^{32} \omega_j \sum_{i \in \Omega_k}^{|\Omega_k|} SDG_{gap}^{i,j}, \tag{5}$$

where Ω_k is the set of indicators belonging to SDG k.

5.3.1. Optimal allocation of fiscal transfers

Fig. 8 presents the results of implementing differential evolution to minimize equation (4) (with the objective function considering all the SDGs). The control variables are the shares of federal transfers across the 32 states. We present the best solution obtained after 200 generations of the algorithm (after 100 generations the solutions are stable). This exercise assumes no solution constraints, other than the shares adding up to 1. Thus, it does not take into consideration rules, established by historical and political factors, that influence the distribution of federal transfers across the states.[24]

Fig. 8 indicates that approximately one-third of the states received, in 2019, a share of fiscal transfers close to the optimal one. This is the case of the state of Chiapas (CHP), whose dot overlaps the dashed line. However, there are notable cases of under-allocations, like the states of Jalisco (JAL) and Veracruz (VER). Cases of over-allocation can be found in the State of Mexico (MEX) and Mexico City (CMX).[25] As mentioned above, these are the states with the strongest tax collection capabilities. Presumably, in an environment with no fiscal rigidities, the model's outcome suggests that it is convenient to reduce the transfers to these

states if the objective is to achieve more equitable development across the 32 entities and all the SDGs. The outcome for Mexico City is particularly salient since the optimal distribution suggests an extremely low share. Perhaps this feature can be explained by the fact that the capital of the country has the largest amount of locally collected taxes, which provides for the required public funding to reduce the SDG gaps.

5.3.2. Emphasizing specific goals

As an alternative optimization objective, a national government may consider a subset of SDGs or a weighted combination of them. This refinement, of course, assumes that, while the federal government is trying to prioritize a specific SDG, subnational authorities still have their own goals in mind, since these governments have different mandates as a result of their local political realities. Whether this is the case in every country is debatable, as some degree of misalignment between federal and subnational authorities should always be expected. To explore if this type of exercise alters the previous results in meaningful ways, we estimate the optimal distribution of federal transfers for the hypothetical cases in which Mexico focuses on a single SDG. This exercise is relevant in the context of Mexico, where there is public debate on the role of the federation as a distributive force to mitigate poverty differences and general income inequality between the states; i.e. there is discussion on whether federal transfers should focus on SDG 1.

In Fig. 9, we display these distributions by SDG. As reference points, we also include the empirical distribution (top ridge with line hatches) and the optimal one that considers all the SDGs (second highest ridge with dot hatches). These ridges facilitate the visual identification of discrepancies in the intergovernmental transfer of federal resources under different fitness landscapes. They also expose salient cases of over and under-allocations. Each one of the SDG-specific ridges is obtained by minimizing equation (5) for the corresponding SDG. Moreover, a high degree of similarity between the empirical ridge and an optimal one indicates that the misalignment of fiscal transfers is small, at least under the chosen objective function and the assumption of an unconstrained reallocation.

The first thing to notice in Fig. 9 is that the distributions of fiscal transfers vary with the focus SDG. That is, there is sensitivity to the type of 'development mode' pursued. When comparing the stripped and dotted ridges, we observe a strong correlation in terms of hills and valleys, but also several states with discrepancies in terms of over and under-allocations. For instance, the high hill for MEX in the empirical ridge is less pronounced than the one obtained from the all-SDGs optimization. Likewise, the optimal distributions for SDG-specific objective functions exhibit distinctive features. For example, the highest allocated transfer does not necessarily correspond to the same state, although three of them repeat as top recipients: MEX when focusing on SDGs 11, 9, 8, 6, 4, and 3; VER for SDGs 16, 10, and 1; and CMX for SDGs 12 and 2. PUE has the leading share in SDG 15, while GUA does in SDG 5. Note that MEX and VER receive a relatively large share of transfers in all the SDG-specific development modes. However, there are several cases like CMX, COL, NLE, SIN, CHP, and GRO that are granted a small share of transfers in some development modes but high in others.

The outcomes presented in this plot cannot be explained by plain intuition. For instance, one would expect that poor states like Chiapas (CHP), Guerrero (GRO), and Oaxaca (OAX) would receive a large share of transfers when the promoted development mode is defined in terms of SDG 1. However, the simulated ridge in the last row shows that this does not always occur. While CHP does receive large fiscal transfers, GRO and OAX do not. In fact, Nuevo Léon (NLE), a relatively affluent state receives a larger transfer share than the latter two. This result can be explained, in part, by the large population of other states (e.g., NLE, MEX, GUA) that attract more resources for ameliorating precarious economic conditions; but most importantly, because states like GRO and OAX can exhibit structural deficiencies that preclude the possibility of abating poverty through sheer increments of public funds. These counter-intuitive results will be better understood once we introduce the

---

[24] In Appendix F, we perform a constrained optimization using similar restrictions to those presented in official documents. This exercise shows that the method can be adapted to the circumstances of different empirical realities.

[25] The linear relationship between the empirical and simulated shares is partly the consequence of population size. Yet, a strong linear relationship does not imply a close similarity in shares since, even in a high correlation setting, there can be numerous cases of over or under-assignments.
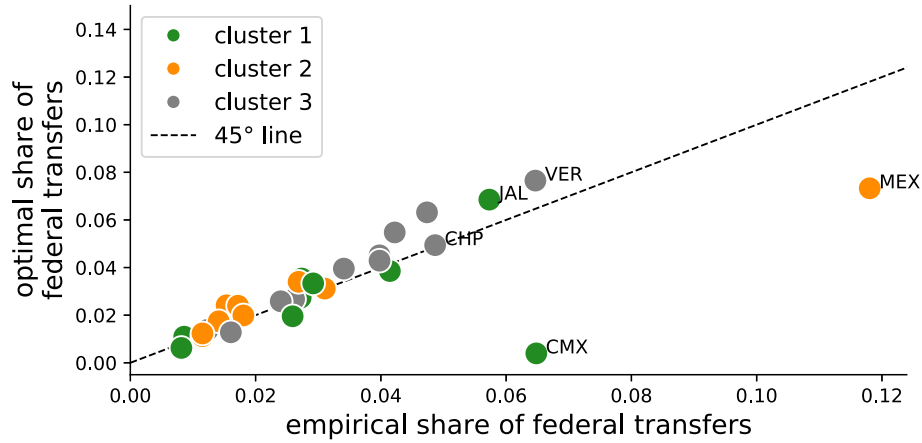
Fig. 8. Empirical versus optimal distribution of federal transfers. *Note*: Each dot corresponds to a state. The top 5 states in terms of received federal transfers (in the data) are presented with their abbreviation. A dot overlapping the 45° line indicates that the empirical federal transfer is very close to the one obtained through the optimization procedure. *Source*: Authors' own calculations.
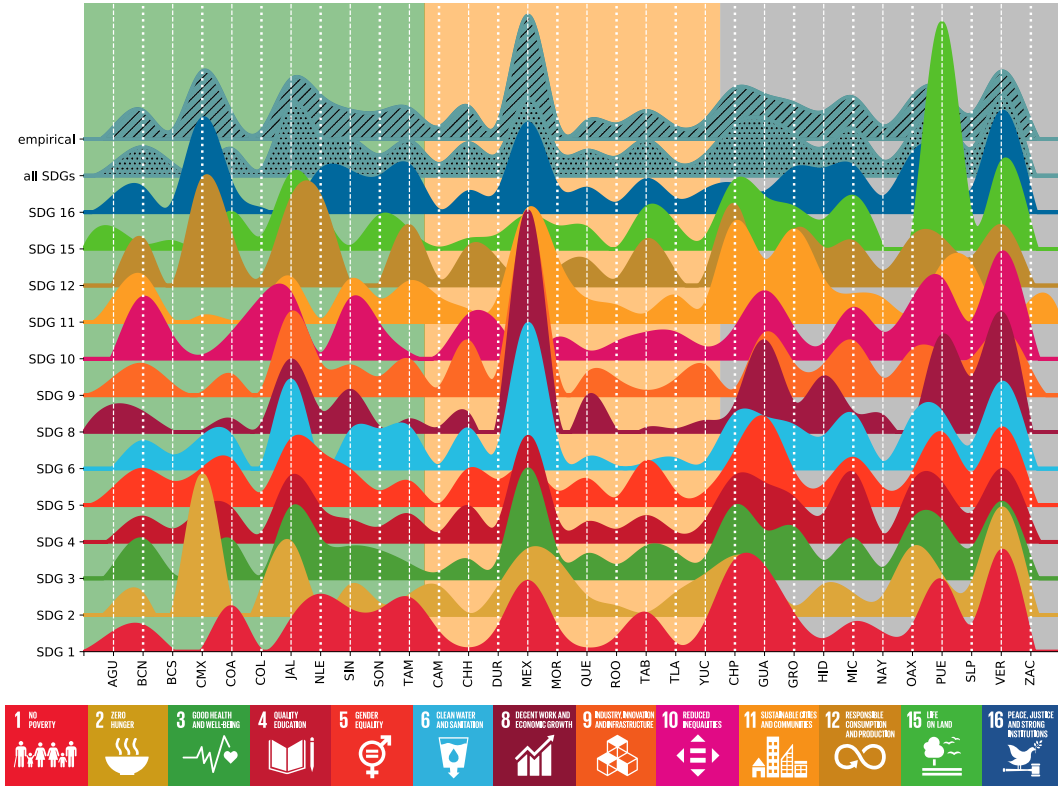
Fig. 9. Optimal reallocation under SDG-specific objective functions. *Note*: The top distribution (stripped) corresponds to the empirical 2019 distribution of federal transfers. The second highest distribution in the plot (dotted) is the optimal allocation of federal transfers when the objective function considers all the SDGs. All other distributions correspond to the optimal allocation of federal transfers when considering only one SDG at a time. The SDG-specific distributions are color-coded according to the logos presented at the bottom of the graphic. The states have been ordered by cluster, as indicated by the background colors. *Source*: Authors' own calculations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

concept of 'budgetary frontiers' in section 5.3.4.

5.3.3. Quantifying similarities between distributions of federal transfers

To provide quantitative insights into the different optimal distributions of federal transfers, we compute the coherence index developed by Guerrero and Castañedañ [22]. This index measures the Euclidean distance that the empirical distribution has with respect to the optimal, and to another one that is the opposite to the optimal (where the smallest allocation goes to the state that received the largest one in the optimal distribution and so forth). If the empirical distribution is closer to the optimal one, the index is positive, denoting a certain level of coherence (or similarity). If it is closer to the anti-optimal one, the index is negative and indicates that the allocation is incoherent. More specifically, the index is computed as

$$coherence = \frac{d(P - P^\circ) - d(P - P^*)}{d(P - P^\circ) + d(P - P^*)}, \tag{6}$$

where P is the empirical distribution, P^* is the optimal one, P° is the anti-optimal one, and $d(\cdot)$ is the Euclidean distance between two vectors. The

index goes from −1 to 1, so a positive unit means that the empirical and the optimal allocations are identical.

We report the coherence index in Table 1, together with additional metrics about gap reductions and time savings. Note that all the index values, except one, are below 0.6, indicating some degree of similarity but not a very high one. The calculated coherence for the all-SDG optimal distribution is 0.571, and it increases to 0.629 when the benchmark ridge is the development mode that focuses on SDG 4. This means that the shares associated with the 'quality of education' criterion are a good reflection of the empirical federal transfers. In contrast, the ridges corresponding to SDGs 2, 8, 10, 11, and 15 are poor reflections of the observed shares in the 2019 budget.

The remaining results in Table 1 are also interesting in so far as they quantify the extent of the improvement that can be attained when using the optimal allocations instead of the empirical one. The gap reductions estimated for the criteria under analysis (all-SDGs and SDG-specific) are not particularly striking. Nevertheless, the average months saved (across states and indicators) in convergence time to the 2030 goals are meaningful. Accordingly, the largest savings can be obtained when fiscal transfers are assigned using the all-SDGs development mode (16.031 months) or the criterion based on indicators of SDG 16 (15.109 months). Conversely, only 4.910 months can be saved when focusing on SDG 8, an interesting outcome that challenges the traditional view of growth-centered development. Finally, the large standard deviations indicate disparities within states and indicators.

5.3.4. Some complexities underpinning the outcomes of decentralization

Table 1
Outcomes from optimal federal transfers.

SDG	Gap reduction (%)	Gap reduction in SDG (%)	Months saved	Coherence
All	0.506 (1.174)	0.506 (1.174)	16.031 (13.779)	0.571
1	0.038 (2.368)	1.192 (2.916)	11.015 (23.765)	0.316
2	0.109 (2.758)	0.605 (1.688)	13.663 (31.806)	0.280
3	0.249 (1.545)	0.474 (0.746)	13.233 (17.090)	0.561
4	0.217 (1.111)	0.506 (0.969)	12.842 (12.456)	0.629
5	−0.128 (1.585)	0.553 (2.267)	9.648 (15.680)	0.447
6	0.152 (1.983)	0.550 (1.034)	12.983 (22.416)	0.538
8	−0.500 (2.815)	0.874 (3.342)	4.910 (28.669)	0.248
9	−0.038 (1.822)	0.559 (2.727)	10.003 (21.540)	0.440
10	−0.228 (2.468)	0.269 (0.459)	9.900 (26.980)	0.086
11	−0.024 (3.309)	2.472 (4.903)	13.635 (38.051)	0.150
12	−0.217 (2.504)	2.189 (3.688)	8.245 (26.090)	0.327
15	−0.222 (2.542)	1.291 (3.051)	9.246 (25.153)	0.268
16	0.357 (2.024)	0.845 (1.696)	15.109 (22.307)	0.441

Note: Averages across states and indicators. The numbers in parentheses correspond to the standard deviation.
Source: Authors' own calculations.

The previous results consider scenarios in which the national government has no restrictions to reallocate a large percent of the shareable tax revenues. Naturally, one may argue that, in the real world, governments face political economy frictions that constrain the space of potential distributions, or the total amount of transfers that they can manipulate. Furthermore, suppose that the federal government would

only be able to allocate freely a subset of these transfers, say, the participations.[26] A linear logic would suggest that the optimal distribution should be proportional to the one found in section 5.3.1 for all transfers, trivializing the redistribution problem into one of re-scaling or shifting an allocation. This, however, is not the case because, as suggested in section 5.2, the sensitivity of the indicators as a response to the size of the budget is non-linear. Thus, if the metric to be optimized relates to SDG gaps, then there is an asymmetric impact between removing resources from well-endowed states and transferring them to not so wealthy ones. In this section, we restrict the distribution of resources to the participations while considering all the SDGs in the objective function.

Panels (a) and (b) in Fig. 10 present the optimal distribution for all the transfers and for the participations only, respectively. From an overview of the two panels, a notorious change in the degree of coherence (i.e., less similarity between the empirical and optimal shares in the second case) is clear, as well as various qualitative differences in terms of states who switch from receiving an over-allocation (under all transfers) to an under-allocation (under participations) such as the State of Mexico (MEX). The observed changes cannot follow from proportional or linear shifts of the optimal distribution estimated for full transfers (since the magnitudes of the over- or under-allocations have different sizes or even directions), but from more fundamental non-linearities related to the political economy of the system. We provide two additional panels that help explaining these intriguing results: panel (c), which contains the empirical budget from Fig. 1, and panel (d), which displays the 'budgetary frontiers' and the average historical gaps closed by each state.

A budgetary frontier is a theoretical concept, introduced by Guerrero and Castañeda [21], that captures the potential of governments to generate significant progress in their indicators through the sheer growth of their budgets, i.e. without reforming long-term structural factors. More specifically, in panel (d), the budgetary frontier is quantified as the SDG gap that would remain in 2030 if the states would have all the needed resources and these were used fully efficiently. To understand why, even with plenty of resources, not all indicators reach their goals, it suffices to examine equations (2) and (1). In order to induce dynamics on the budgetary frontier we set $\gamma_{i,1} = 1$ for every indicator i and period t. In this way, the speed of convergence is fully determined by the long-term factors captured in α_i (from equation (1)), which are calibrated using the empirical data.

As an example, we can focus on the state of Querétaro (QUE), which has been one of the most successful development cases in the country during the past decade. This state shows the smallest SDG gap when working on the budgetary frontier (lowest bar in panel (d)), while it also exhibits important progress in the historical advancement of its indicators (black marker in panel (d)). From panel (c), we can see that QUE received substantially less budget than other states with similar progress, which means that every peso spent in that entity was substantially more productive than in any other state–hence the significantly lower SDG gap in panel (d). Consequently, due to its structural factors, QUE can achieve more through sheer expenditure.

Now that we are equipped with the intuition behind budgetary frontiers, we can interpret the differences between panels (a) and (b). For this, we focus on three states and analyze their allocations under the two optimal distributions. First, let us concentrate on the State of Mexico (MEX). When optimizing for all transfers (panel (a)), the solution indicates that it should receive less financing. However, when focusing on the participations (panel (b)), the result suggests that it should receive more than what it received in 2019. To understand this, notice that,

---

[26] Restricting a redistribution to the participations is a realistic setting for Mexico. The remaining transfers, the contributions, are usually tied to specific topics such as education or public health, so one could expect certain rigidity when it comes to reallocating those resources.
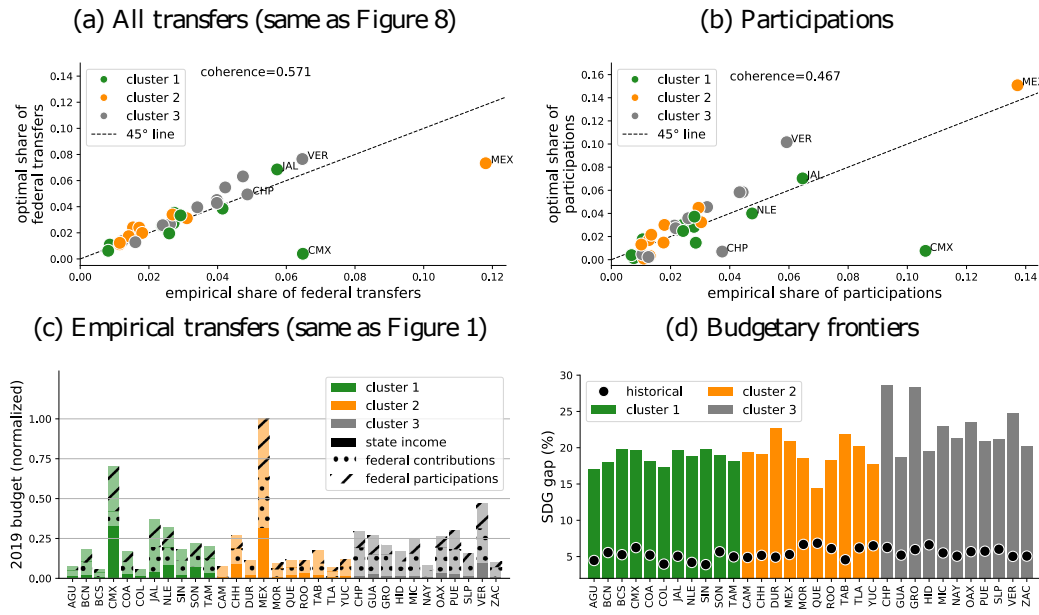
Fig. 10. Optimal distribution of participations versus all federal transfers. *Note*: The black markers in panel (d) indicate the average gap closed by each state during the sampling period (2005–2018) in percentage. *Source*: Authors' own calculations and INEGI (inegi.org.mx/programas/finanzas) for panel (c).

when optimizing for all transfers, several states can become extremely vulnerable because most of their budget comes from this source of income. In fact, the convergence time of an indicator approaches infinity when the budget tends to zero due to equation (2) (hence the asymmetric shape of the sensitivity curves presented in Fig. 7). In contrast, MEX has substantial self-generated resources, so the costs (in SDG gaps) produced by removing transfers are less than the benefits from assigning them to other states. This lack of resources among states with weak fiscal capacities diminishes when focusing only on participations (see panel (c)). In this case, the costs of transferring resources from MEX to states from cluster 3 outweigh the benefits. In fact, since MEX presents better performance on the budgetary frontier than most states from cluster 3, the optimization algorithm finds that there are more potential gains in allocating further participations to MEX.

Next, let us look at Mexico City (CMX), which has the second-largest budget but that, in contrast to MEX, is recommended to receive fewer transfers in both exercises. The explanation of why the optimal allocation share for CMX does not increase from panel (a) to (b) has to do with the returns to expenditure captured in parameter β. MEX has almost twice the population of CMX, but their difference in budgets is not that large. This means that every peso invested in CMX is more productive than in MEX (CMX has a larger β, see Appendix D), so this effect can compensate for reductions in CMX's transfers. In fact, CMX performs better than MEX on the budgetary frontier (see panel (d)), and shows one of the best historical performances in the country (substantially better than MEX). This implies that CMX's higher returns help to mitigate the impact of budgetary reductions, while potential increments would yield marginal returns that are no better than those obtained from transferring participations to other states, reason why the optimization procedure maintains CMX as an over-allocated entity.

Finally, let us look at the state of Chiapas (CHP), a historically marginalized region of Mexico and the least developed state in the country according to the HDI. Why would the optimization algorithm suggest that CHP is currently receiving the right amount of transfers (panel (a)) and that it should receive fewer participations when optimizing for the latter, especially given its deterred socioeconomic and environmental situation? The answer lies in CHP's performance on the budgetary frontier, which is the worst across all states. In a setting where all transfers can be reallocated for finding the optimal shares, CHP would be severely hit by a large removal of resources.

However, given a budgetary baseline through the contributions, the result suggests that participations are not necessary, as its performance would be quite poor.[27] Instead of additional federal transfers, the type of intervention that CHP needs is one of a structural, long-term, nature that can push its budgetary frontier (parameter α). Likewise, assigning participations to CHP would translate into missed opportunities in other states. In fact, these opportunities seem to be capitalized by Veracruz (VER), a cluster-3 state with better performance on the budgetary frontier. What stands out of this line of reasoning is the fact that the identification of such opportunities depends critically on the overall configuration of allocations and budgetary frontiers, as well as their interdependencies. In other words, finding an optimal distribution of transfers is not a trivial task that can be undertaken by assuming states as homogeneous silos since a government cannot increase resources into a state without causing asymmetrical disturbances in others.

6. Discussion and conclusion

Through a computational approach, we open a new window for the study of regional development and fiscal federalism from the perspective of public spending. The models used for these analyses, and others, should be sophisticated enough to describe the key elements behind the complexity of the phenomenon under study. This paper is a first attempt in this direction, focusing on the Mexican case. The associated agent-based model is calibrated with a comprehensive dataset of SDG indicators and budgetary information across 32 states.

This computational tool is designed to analyze budgetary allocations within a set of government programs. In this paper, we present an extension to this tool for studying budgetary allocations within subnational governments. Although the model does not address explicitly all the intricacies that exist across government entities, either horizontally–across policy issues–or vertically–across government levels, its main features produce simulation outputs that have been previously

---

[27] There seems to exist critical points (in the allocation of transfers) at which a state's development collapses, but after which one cannot achieve substantially larger progress. Importantly, this critical point is potentially linked to various factors included the quality of governance, long-term structural factors, and spillover effects.

validated [21–24] with comprehensive datasets, providing a theoretical and empirical foundation for the study of vertical intergovernmental transfers and their implications for sustainable development.

Discovering the optimal allocation of federal transfers across subnational units is far from being a task that can be solved through traditional *ceteris paribus* back-of-the-envelope exercises (e.g., a correlation between transfer shares and a poverty index). Instead, one needs to consider–simultaneously–factors such as the stage of development, expenditure returns, budgetary frontiers, and how they interact within and between different regions. Only when these features are taken into consideration it is possible to measure statistical regularities behind a development process, for instance, the opportunity cost of transferring fiscal resources to a particular state in terms of the non-realized benefits of developing others. This is why the adoption of a systemic view, the acknowledgment and formalization of complex dynamics, and the deployment of computational tools are so important for tackling this kind of problems.

As a summary of the most important results of our study, we mention the following. (i) We find that most of the SDG indicators will attain levels below their goals in 2030, in spite of the fact that these were established with reasonable criteria. This outcome indicates that long-term structural factors, beyond the budget and how it is allocated, need to be addressed by policymakers. (ii) We estimate the non-linear response of the indicators to the budget size, and find that it varies between indicators and clusters. Thus, modifications in the distribution of federal transfers have consequences for the co-evolution of indicators between and within states. (iii) We identify over- and under-allocations of federal transfers across specific states, noting that these inefficiencies may vary with the specific objective function that describes the nation's aspirations. For example, potential objectives may include having a wide set of goals like the whole range of SDGs, or focusing on a specific topic such as poverty or public health. (iv) With our coherence index, we show that some objective functions, when used as development modes (e.g., 'quality of education'), are more accurate than others (e.g., 'reduced inequalities' and 'sustainable cities and communities') in

describing empirical budget shares. (v) We also find that, if a comprehensive objective function were to be used to allocate transfers, there could be substantial savings in terms of convergence time to reach the 2030 Agenda.

Finally, we would like to emphasize that in this paper we study problems associated with fiscal federalism from the spending side. Although we are aware that important questions to be solved also come from the revenue side. One of the shortcomings of our model is, precisely, that we do not introduce any consideration associated with taxation issues at the subnational level. Therefore, one natural extension of the model is to establish certain connections between development indicators and the central government's overall budget; including possible trade-offs in which more taxation might imply possible setbacks in the evolution of some indicators (e.g., GDP). However, we believe that, before incorporating more theoretical subtleties in the model, it is important to have richer databases; otherwise, there would be several free parameters that could not be calibrated properly.

CRediT authorship contribution statement

Omar A. Guerrero: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Gonzalo Castañeda:** Conceptualization, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Georgina Trujillo:** Data curation, Investigation, Writing – review & editing. **Lucy Hackett:** Data curation, Investigation, Writing – review & editing. **Florian Chávez-Juárez:** Conceptualization, Data curation, Investigation, Project administration, Writing – review & editing.

Acknowledgement

A Data.

A.1 SDG indicators

Table A1 gives some summary statistics of the dataset. In Panel A, we report the total number of observations and missing values across states and indicators within each SDG. In Panel B, we show the indicator minimum, maximum, and average number of observations (original, not imputed) within a given SDG (at the indicator-state level). Several SDGs have indicators with as few as 3 observations; these correspond to census-based indicators collected every 5 or 10 years. However, the overall average number of observations is almost 10. This implies that, on average, each indicator has observations for well over half of the years in the sampling period. In Panel C, we specify the number of indicators provided by the three most prominent sources, plus the remaining ones: 'other govt.'.[28] The 'other govt.' category is comprised of 30 sources from varying surveys and dependencies of the Mexican government such as the Tourism Ministry, the National Water Council, the Executive Secretary of the National Public Security System, etc. The four World Bank indicators come from the 'Doing Business' database, which is provided at the subnational level by the source.

For a more disaggregated summary of the data, Table A2 lists each indicator, the SDG it is assigned to, and the method used for imputing the numeric goal for that indicator, as described in Section 4.1. Finally, Figure A1 gives a visual representation of how average indicator levels vary by state.

Table A.1
Data summary

A. Observations														
SDG	1	2	3	4	5	6	8	9	10	11	12	15	16	All
Total no. obs.	31	6	137	125	77	57	154	112	15	30	9	54	217	1024
Total no. missing	53	8	45	15	8	42	45	39	41	12	5	19	119	451

---

[28] CONEVAL is the National Council for the Evaluation of Social Development Policy, a decentralized dependency of the Mexican government that measures poverty and inequality. INEGI is the National Institute for Statistics and Geography, which administers surveys, as well as the census, and produces the national accounts.

(*continued*)

A. Observations

SDG	1	2	3	4	5	6	8	9	10	11	12	15	16	All
Total no. obs.	31	6	137	125	77	57	154	112	15	30	9	54	217	1024
Total no. missing	53	8	45	15	8	42	45	39	41	12	5	19	119	451

B. Observations by indicator

SDG	1	2	3	4	5	6	8	9	10	11	12	15	16	All
min	3	6	6	6	8	3	4	9	3	6	9	10	3	3
max	6	6	14	14	14	13	14	14	6	13	9	12	13	14
mean	5.17	6.00	10.54	12.50	12.83	8.14	11.00	11.20	3.75	10.00	9.00	10.80	9.43	9.94
std	1.17	0.00	2.07	2.59	2.40	4.18	4.33	2.04	1.50	3.61	0.00	1.10	3.30	3.57
SDG	1	2	3	4	5	6	8	9	10	11	12	15	16	All
min	3	6	6	6	8	3	4	9	3	6	9	10	3	3
max	6	6	14	14	14	13	14	14	6	13	9	12	13	14
mean	5.17	6.00	10.54	12.50	12.83	8.14	11.00	11.20	3.75	10.00	9.00	10.80	9.43	9.94
std	1.17	0.00	2.07	2.59	2.40	4.18	4.33	2.04	1.50	3.61	0.00	1.10	3.30	3.57

C. Indicators by source

SDG	1	2	3	4	5	6	8	9	10	11	12	15	16	All
CONEVAL	3	1	1	1	0	0	0	0	1	1	0	0	0	8
INEGI	2	0	2	2	4	4	4	0	3	1	0	0	2	24
Other govt.	1	0	10	7	2	3	6	10	0	1	1	5	21	67
World Bank	0	0	0	0	0	0	4	0	0	0	0	0	0	4
Total	6	1	13	10	6	7	14	10	4	3	1	5	23	103
SDG	1	2	3	4	5	6	8	9	10	11	12	15	16	All
CONEVAL	3	1	1	1	0	0	0	0	1	1	0	0	0	8
INEGI	2	0	2	2	4	4	4	0	3	1	0	0	2	24
Other govt.	1	0	10	7	2	3	6	10	0	1	1	5	21	67
World Bank	0	0	0	0	0	0	4	0	0	0	0	0	0	4
Total	6	1	13	10	6	7	14	10	4	3	1	5	23	103

Note: Panel A counts all observations and missing observations per state, summing across indicators and years. Panel B provides details on the number of observations per indicator and state. Panel C is also at the indicator-state level and counts the number of indicators broken down by source.
Source: Authors' own calculations.

Table A.2
Indicators and their goals

SDG	Indicator	Indicator abbrev.	Method
1	Percent of the population in extreme poverty	pct_extreme_pov	1
1	Percent of the population vulnerable because of low income	pct_vulner_inc	5
1	Percent of homes with some level of overcrowding	pct_overcrowd	5
1	Human Development Index	human_dev_index	5
1	Percent of young people aged 19–29 with income below the welfare line	pct_below_welfare	OECD
1	People who left poverty	pop_left_pov	OECD
2	Percent of the population lacking adequate access to food	pct_hungry	1
3	Mortality rate of diabetes mellita per 100,000 inhabitants	diabetes_mort	1
3	Fertility rate of women aged 15-19	teen_fert	1
3	Percent of the population lacking adequate access to health services	pct_no_healthcare	1
3	Proportion of 1-year-old infants with a full basic vaccination record	pct_full_vaccine	2
3	Mortality rate of HIV/AIDS (per 100,000 inhabitants)	aids_mort	2
3	Maternal mortality (deaths per 100,000 live births, estimated)	maternal_mort	5
3	Proportion of births attended by a trained medical professional	births_with_doc	5
3	Nurses in public health institutions per 1,000 inhabitants	nurses_per_pop	5
3	Infant mortality rate	infant_mort	OECD
3	Doctors in public health institutions that have contact with patients per 1,000 inhabitants	doctors_per_pop	OECD
3	Licensed hospital beds per 100,000 inhabitants	hops_beds_rate	OECD
3	Life expectancy at birth	life_expect	OECD
3	Neonatal mortality rate	neonat_mort	OECD
4	Net enrollment rate in middle school (12–14 years of age)	enroll_secondary	2
4	Net enrollment rate in pre-school education (2–5 years of age)	enroll_preschool	2
4	Absorption rate in high school	absorp_high_school	2
4	Absorption rate in undergraduate education	absorp_college	2
4	Literacy rate in young adults (15–24 years old)	pct_literate	2
4	Percent of the population aged 16+ or born after 1982 with educational deficiencies	pct_educ_deficient	2
4	Terminal efficiency in high school	finish_high_school	5
4	Museums per 100,000 inhabitants	museums_rate	5
4	Libraries per 100,000 inhabitants	library_rate	5
4	Proportion of the labor force with high school education or more	pct_workers_educated	OECD
5	Proportion of institutions within the organizational structure of public administration headed by women	pct_women_govt	1
5	Ratio men-women in the National System of Researchers (SNI)	gender_eq_research	1
5	Annual brute rate of deaths due to homicides of female victims	fem_homicide_rate	5
5	Percent of working mothers aged 15+ with access to childcare	access_childcare	5

(*continued*)

| SDG	Indicator	Indicator abbrev.	Method
5	Proportion of non-agricultural workers aged 15 and up who are women	pct_women_non-ag	OECD
5	Ratio men-women aged 15+ in the economically active population	gener_eq_workers	OECD
6	Proportion of the population with access to running water	pct_running_water	2
6	Population with access to sewage and basic sanitary services	pct_access_sewage	2
6	Percent coverage in treatment of waste waters	waste_water_treat	5
6	Surface water quality index	surf_water_quality	5
6	Operational water treatment capacity	water_treat_capacity	5
6	Percent of homes with water provided by truck	pct_water_by_truck	5
6	Spending on operations in industrial waste water treatment plants	spend_water_treat	5
8	GDP per capita	gdp_per_cap	5
8	Rate of informal employment	inform_emp_rate	5
8	Rate of underemployment	underemp_rate	5
8	Unemployment rate of young people aged 15-29	unemp_rate_young	5
8	International Commerce	intl_commerce	5
8	Direct foreign investment	direct_foreight_invest	5
8	Agricultural productivity (Value/hectare)	ag_productivity	5
8	Labor prductivity index, based on the employed labor force	labor_prod	5
8	Workers enrolled in IMSS as a percent of the total population	pct_enrolled_IMSS	5
8	Construction permit cost index (% of income per capita)	construc_cost_index	5
8	Costs of contract enforcement index (% of income per capita)	contract_cost_index	5
8	Property registery costs index (% of income per capita)	prop_reg_cost_index	5
8	Business start-up costs index (% of income per capita)	business_cost_index	5
8	Net employment rate for adults aged 15 and up	emp_rate	OECD
9	Hotel nights per capita	hotel_stays_per_cap	1
9	Number of commercial banks and ATMs per 100,000 inhabitants	bank_atm_rate	5
9	Government-registered researchers per 100,000 economically active adults	researchers_rate	5
9	Budget assigned to science and technology in mixed state funds	budget_science	5
9	Natural disaster resiliance index	disaster_resil_index	5
9	Energy intensity of the economy (MWh per 1,000,000 GDP per year)	energy_intensity	5
9	Debit cards per 10,000 inhabitatnts	debit_card_prev	5
9	Credit cards per 10,000 inhabitants	credit_card_prev	5
9	Paved highways as a percent of all highways	pct_highway_paved	5
9	Number of patents solicited per 1,000,000 inhabitants	patents_rate	OECD
10	Index of equity in access to health services	equity_in_health	2
10	Index of educational equity in middle school	gender_equity_second	5
10	Index of equity in access to drainage	equity_in_drainage	5
10	State Gini coefficient	gini	OECD
11	Percent of the population lacking adequate access to basic services to their home	pct_no_basic_serv	1
11	Property registration	property_reg	5
11	Volume of garbage and waste generated (kg. per person)	vol_garbage	5
12	Businesses certified as green per 1,000 businesses	green_business_rate	5
15	Forest fires: affected surface area (pct. of total forest surface)	forest_fires	5
15	Budget assigned to the CONAFOR national program for forests as a proportion of the area supported	budget_forest_dev	5
15	Vegetation planted per km. square of reforested areas	reforest_plants	5
15	Reforested area as a pct. of forest cover	pct_area_reforest	5
15	Area supported by the Forest Development Program as pct. of forest	forest_dev_program	OECD
16	Participation of the eligible population to vote in federal elections	pct_particip_elec	3
16	Rate of resolution in the penal justice system	penal_resol_rate	5
16	Debt service as a proportion of total income	pct_debt_serv	5
16	Judges per 100,000 inhabitants	judges_per_pop	5
16	Number of public prosecutors per 100,000 inhabitants	prosecute_per_pop	5
16	Number of computers per 100 public servants in state public administration	computers_in_govt	5
16	Extortion rate	extortion_rate	5
16	Home burglary rate	burglary_rate	5
16	Rate of commercial burglary	commerc_burg_rate	5
16	Rate of vehicle theft with or without violence	vehicle_theft_rate	5
16	Kidnapping rate	kidnapping_rate	5
16	Financial autonomy	financial_auton	5
16	Investment capacity (percent of govt. spending on investment)	invest_capacity	5
16	Government income as a percent of GDP	govt_income	5
16	Index of transparency and availability of state fiscal information	fiscal_transp_index	5
16	Total unpayed to the federal contribution fund	unpayed_contrib	5
16	Percent of the population who has been a victim of corruption in at least one government process	pct_victim_corrup	5
16	Total progress in budget and impact evaluation (PbR-SED)	budget_prog_index	5
16	Political system sub-index	political_sys_index	5
16	State budget information index	budget_info_index	5
16	Crime rate per 100,000 inhabitants ages 18-29	crime_rate	OECD
16	Rate of registered intentional homicide	intent_homicide_rate	OECD
16	Perception of corruption in the federal government	perceived_corrup	OECD
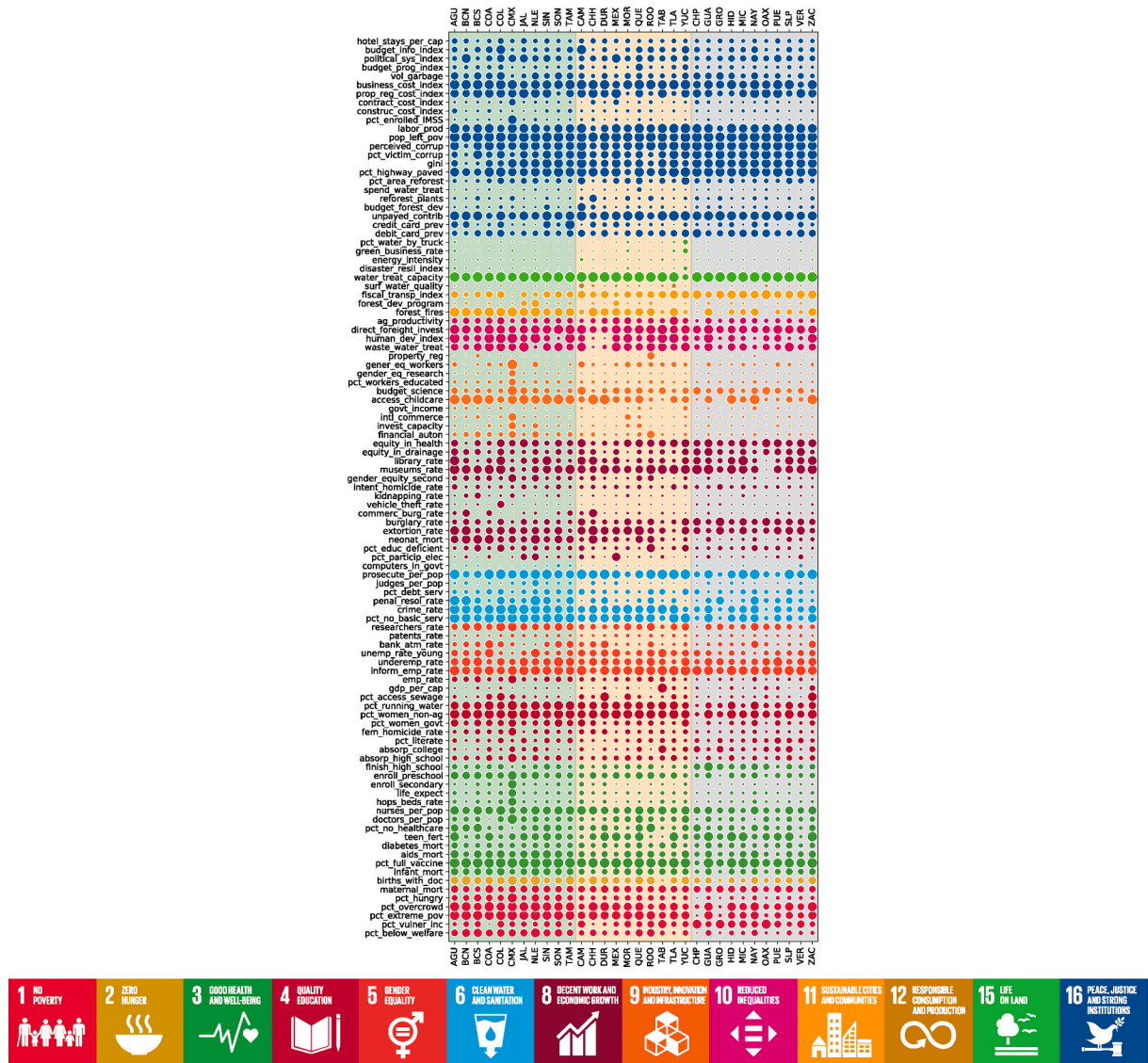
Figure A.1
Average indicator levels by state. *Note*: Each dot size is proportional to the level of the indicator. The background colors correspond to the three clusters. *Source*: Authors' own calculations.

A.2 Budgetary data

In Mexico, INEGI generates annual information on the public finances of states and municipalities. Data on public income and expenses of the states are generated from administrative records and key informants from core state agencies such as the Ministry of Finance, State Administration, or Oversight Bodies. Information on both income and expenses is harmonized by INEGI across 11 categories of income and 11 more of expenditure. The former include gross total revenue, total net revenue (gross income minus an initial availability that represents assets), and revenue disaggregated by source. It also differentiates, states-collected revenue from federal transfers coming from shared funds (each fund is regulated by specific sections of the Fiscal Coordination Act).

A.3 Normalization and imputation

Here, we provide a brief description of the procedure of normalization and imputation from Guerrero and Castañeda [21]; as well as a brief discussion of the benefits of this procedure. First, we normalize the indicators to be in the range [0,1]. The idea behind this step is to obtain roughly the same magnitudes across all indicators, and this can be achieved by taking rates, or by normalizing using the feasible limits of an indicator's range. While this normalization is not strictly necessary, it is helpful for the calibration procedure because it reduces the search space of the parameters α; something quite common in machine learning algorithms.[29]

---

[29] For an easier interpretation, the higher values of an indicator denotes better outcomes. Since we normalized the indicators, we can reverse them by using the complement: $1 - I_{i,t}$.

Once the raw data have been normalized, we proceed to impute missing values using the Multi-Output Gaussian Process Toolkit (MOGPTK) [33]. This method employs Gaussian processes and neural networks to predict missing values while exploiting the observations of other 'similar' units of analysis of the same indicator. To define a set of states that are similar to a given state k, we take all the entities that are geographical neighbors of k. Note that, because no two states have the same neighbors, their imputations are unique, which helps to maintain context-specificity. While the MOGPTK represents the state of the art in data-imputation methods, there is still the possibility of extrapolations that lie well beyond the average behavior of an indicator in a given state (for example, if the neighbors exhibit higher volatility). Thus, the imputed data are subjected to a variance correction procedure described in Guerrero and Castañeda [21]. This method preserves the ordinal quality of the extrapolations but compresses its variance to match the empirical data.

B Model details

This appendix is a transcription of the model explanation provided in Guerrero and Castañeda [21]. This explanation spares the reader from interpretations, motivations, and references to the literature, since these can be found in Guerrero and Castañeda [22,23].

B.1 Policy-making agents

There are n agents (or public officials), each in charge of a public policy that is specific to a single policy issue. To implement the mandated policy in a given period t, agent i receives $P_{i,t}$ resources from the central authority (or government). With these resources, the public official tries to leverage two potential benefits: (1) the reputation from being a proficient public servant and (2) the utility derived from being inefficient according to

$$F_{i,t+1} = \Delta I_{i,t}^* \frac{C_{i,t}}{P_{i,t}} + (1 - \theta_{i,t}\tau) \frac{(P_{i,t} - C_{i,t})}{P_{i,t}}, \tag{7}$$

where $F_{i,t+1}$ represents the benefit or utility obtained in the next period. The first summand in equation (7) captures the benefit of being proficient. $\Delta I_{i,t}^*$ is the change in indicator i with respect to the previous period (its performance), relative to the changes of all other indicators. More specifically, the relative change in indicator i is computed as

$$\Delta I_{i,t}^* = \frac{I_{i,t} - I_{i,t-1}}{\sum_j I_{j,t} - I_{j,t-1}}, \tag{8}$$

and it captures the idea that the central authority compares and evaluates the relative performance of each public official, and their implemented policies, through the corresponding development indicators.

Going back to the first summand in equation (7), we find that the relative change in the indicator is pondered by $\frac{C_{i,t}}{P_{i,t}}$. Here, $C_{i,t}$ is the fraction of the allocated resources $P_{i,t}$ that are effectively used towards the policy. We call it the *contribution* of agent i.

Next, let us focus on the second addend of equation (7), which corresponds to the utility derived from being inefficient. Here, $P_{i,t} - C_{i,t}$ is the benefit extracted from not devoting resources to the policy. Thus, when dividing by $P_{i,t}$, it represents the level of inefficiency. Monitoring and penalties may hinder inefficiencies. This is captured by factor $(1 - \theta_{i,t}\tau)$. Variable $\theta_{i,t}$ is the binary outcome of monitoring inefficiencies. If $\theta_{i,t} = 1$, it means that the government has spotted agent i in inefficient behavior. In that case, i is penalized by a factor τ, such that the benefit from these private gains are reduced.

To model the binary outcomes of monitoring efforts, we assume that, every period, an independent realization of $\theta_{i,t}$ takes place for each indicator. This is nothing else than a Bernoulli process with a probability of success $\lambda_{i,t}$ determined by

$$\lambda_{i,t} = \varphi \frac{P_{i,t} - C_{i,t}}{P_t^*}, \tag{9}$$

where P_t^* is the largest allocation in period t. Parameter φ in equation (9) corresponds to the quality of the monitoring efforts.

If an agent becomes more inefficient and their benefits increase, then reinforcement learning takes place, becoming more inefficient the next period. If, in contrast, the government is able to penalize, according to the learning process, they become more proficient the next period. Formally, action $X_{i,t}$ of agent i can be modeled as

$$X_{i,t+1} = X_{i,t} + \mathrm{sgn}((X_{i,t} - X_{i,t-1})(F_{i,t} - F_{i,t-1}))|F_{i,t} - F_{i,t-1}|, \tag{10}$$

where $\mathrm{sgn}(\cdot)$ is the sign function. In order to map action $X_{i,t}$ into the value of the effective resources, we define

$$C_{i,t} = \frac{P_{i,t}}{1 + e^{-X_{i,t}}}. \tag{11}$$

B.2 The government agent

Policy priorities are represented by the allocation profile $P = P_1, \dots P_n$. It is important to introduce a distinction between those indicators that can be intervened via public policies: *instrumental*; and those that cannot: *collateral*. An instrumental indicator exists if the government has a policy or program to directly impact it (i.e., it receives resources). In contrast, a collateral indicator cannot be directly impacted; it is a composite aggregation of various topics, for example, GDP per capita or financial development. Policy priorities can only be defined on the n instrumental indicators, while there can only be n public officials (one in charge of each instrumental indicator). When talking about all the indicators together, we say that there are

$N \geq n$ policy issues in total.

The objective of the government is to close the gap between the goals and the indicators by solving the problem

$$\min\left[\sum_i^N (T_i - I_{i,t})^2\right] \tag{12}$$

through the allocation of budgetary resources across policy issues. The central authority achieves this by adapting its allocation profile P.

In the real world, identifying the precise mechanisms through which governments establish their budgets is extremely challenging. A starting point is the principle of 'gaping', which suggests that governments prioritize the most laggard topics as these may be development bottlenecks. Nevertheless, the political process also introduces adaptations motivated from signals such as the people's demands, and the performance of the different expenditure programs. In the political science literature, these budgetary changes exhibit punctuated dynamics and are modeled through simple stochastic processes [34]. Thus, we combine all these insights in a government heuristic where the policy priorities are established according to

$$P_{i,t} = B \frac{q_{i,t}}{\sum_j q_{j,t}}, \tag{13}$$

where $q_{i,t}$ is the propensity to spend in policy issue i in time t, and B is the budget available in time t.

The evolution of the policy priorities takes place through the propensities. In the first period, these are determined by the normalized gaps

$$q_{i,0} = \frac{T_i - I_{i,0}}{\max(T_. - I_{.,0})}. \tag{14}$$

Then, as time progresses, the propensities are updated according to

$$q_{i,t} = q_{i,t-1} + U(0,1)\left(\sum_k^{t-1} \theta_{i,k}\right)^{-1} \sum_{k|\theta_{i,k}=1}^{t-1} \frac{P_{i,k} - C_{i,k}}{P_{i,k}}. \tag{15}$$

The previous equation is rather intuitive. The term $U(0, 1)$ is a random draw from a uniform distribution in the (0,1) interval. This captures the randomness of societal signals received by the government (it is consistent with the stochastic processes used to model budgetary changes in the literature). The remaining terms to the right correspond to the inter-temporal average inefficiency, which lies in the interval [0,1]. Therefore, the government encourages increments among the most efficient policymaking agents. Note that, in general, the contribution $C_{i,t}$ is not observable by the government, unless there is a successful audit by the monitoring authority. This is why equation (15) conditions the efficiency bias in the allocation of the budget to successful outcomes of the monitoring random variable $\theta_{i,t}$. Thus, the government tends to be more inquisitive with policymakers whose inefficiencies have been spotted in the past.

B.3 Indicator dynamics

We model indicator dynamics through a random growth process. Let γ_i denote a probability associated with the growth process experienced by indicator i. This probability depends on a combination of network effects (i.e., incoming spillovers) and budgetary allocations. Therefore, the growth process is modeled as independent Bernoulli trials with a probability of success

$$\gamma_{i,t} = \beta \frac{C_{i,t} + \frac{1}{n}\sum_j C_{j,t}}{1 + e^{-S_{i,t}}}, \tag{16}$$

where β is a normalizing parameter and $S_{i,t}$ are the net amount of spillovers received by indicator i in time t (this could be positive or negative). The spillovers are computed every period according to $S_{i,t} = \sum_j \mathbf{1}_{j,t} \mathbb{A}_{j,i}$, where $\mathbf{1}$ is the indicator function: 1 if indicator j grew in the previous period and 0 otherwise.

Next, we define the difference equation of indicator i as

$$I_{i,t+1} = I_{i,t} + \alpha_i \xi(\gamma_{i,t}) \tag{17}$$

where $\xi(\cdot)$ is the binary outcome (0 or 1) of a growth trial.

C SDG networks and their estimation

The adjacency matrix ($\mathbb{A}_{j,i}$) in equation (2) captures the structure of interdependencies between SDG indicators (policy issues) and, hence, when estimated with historical data reflects relationships that cannot be modified in the short-term. In our model, an SDG (or spillover) network is considered an exogenous input that describes the context that prevails in a state. Consequently, for a large set of indicators, historical time series are needed to produce state-specific adjacency matrices. Ospina-Forero et al. [28] provide a comprehensive review of quantitative methods for estimating SDG networks. From that study, we conclude that the Bayesian approach of Sparse Gaussian Bayesian Networks developed by Aragam et al. [29] (and known as `sparsebn`) works well when the extension of the series available is relatively small, despite dealing with a high-dimensional policy space, and reduces the possibility of obtaining false positives links (hence the "sparse" term in the name).

The `sparsebn` method estimates a structural equation model and returns a weighted directed network of conditional dependencies where the edges have been filtered in order to minimize potential overfitting (hence the sparseness of the topology). We apply it on to the first differences of the indicators' time series (for the period 2006–2016). To improve the estimation, we pool the observations of the neighboring states of each state, for each indicator, allowing us to build larger datasets. This is recommended to eliminate potential false positives due to inertial trends that are attributable to the natural temporal evolution of the data (and not to their interdependencies). Next, to improve the estimation, those edges

considered–by expert knowledge–to be false positives are removed. Finally, to prevent distorting effects due to extreme values in edge weights (outliers), the magnitude of the maximum and minimum weights is bound by the 95th percentile of the weights' magnitude distribution (the distribution of the absolute values of the weights).

It is important to recall that these networks should not be interpreted as causal relations, but as conditional probabilities. This means that a link $A \rightarrow B$ does not imply that ΔA guarantees ΔB. Accordingly, in this setting, a spillover affects the probability of success γ_i of a public policy in the model, but not necessarily the magnitude of the outcome. Although sparsebn assumes that no temporal dependence exists between observations, we reduce this possibility in the dataset by computing the first differences of the series. Moreover, the extension of the series is enlarged by clustering information of each particular state with that of its three closest geographical neighbors–that according to the data show structural affinities. This method produces state-specific networks, as opposed to other approaches using pooled data. In this fashion, we allow the topology of each state's network to exert an influence on the simulation's outcomes. Castañeda et al. [24] and Guerrero and Castañeda [22,23] show that removing the spillovers from the analysis results in a lower variation of inefficiency across policy issues.
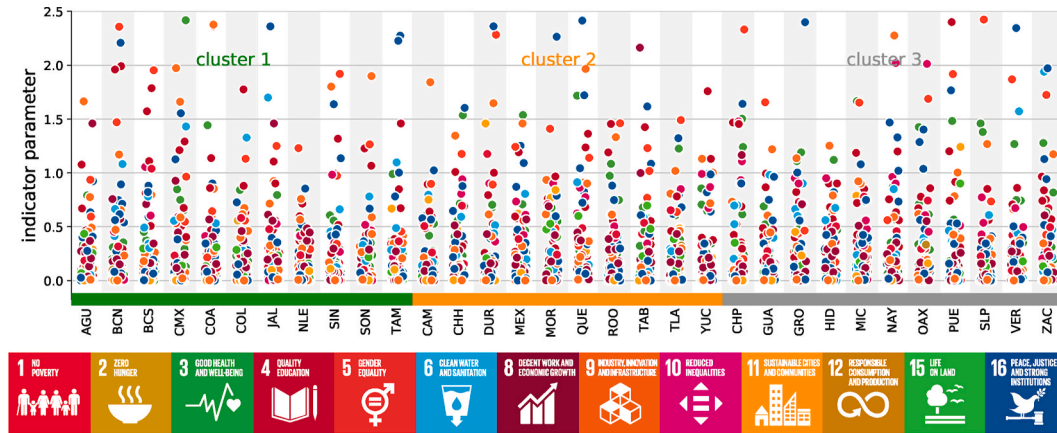
D Parameters

Figure D.1
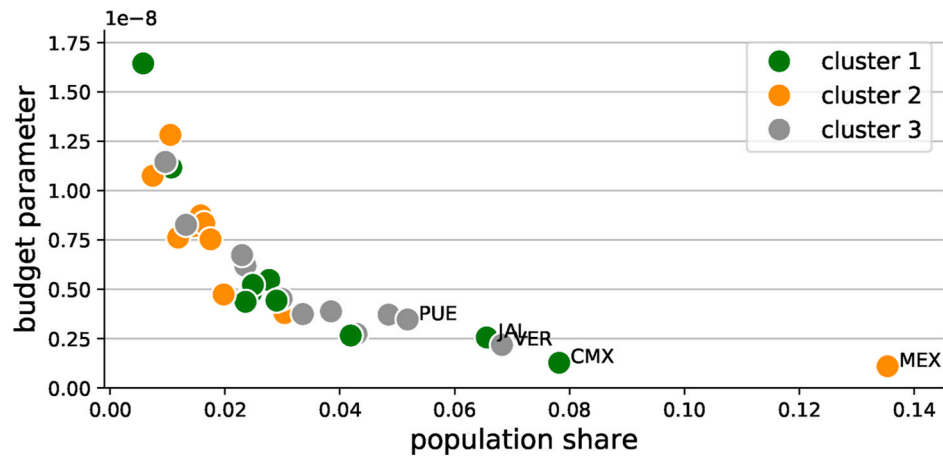Calibrated growth factors (indicator parameters). *Source*: Authors' own calculations.

Figure D.2
Calibrated budget parameters. *Source*: Authors' own calculations.

E Development gaps

ARTICLE IN PRESS

O.A. Guerrero et al.   Socio-Economic Planning Sciences xxx (xxxx) xxx
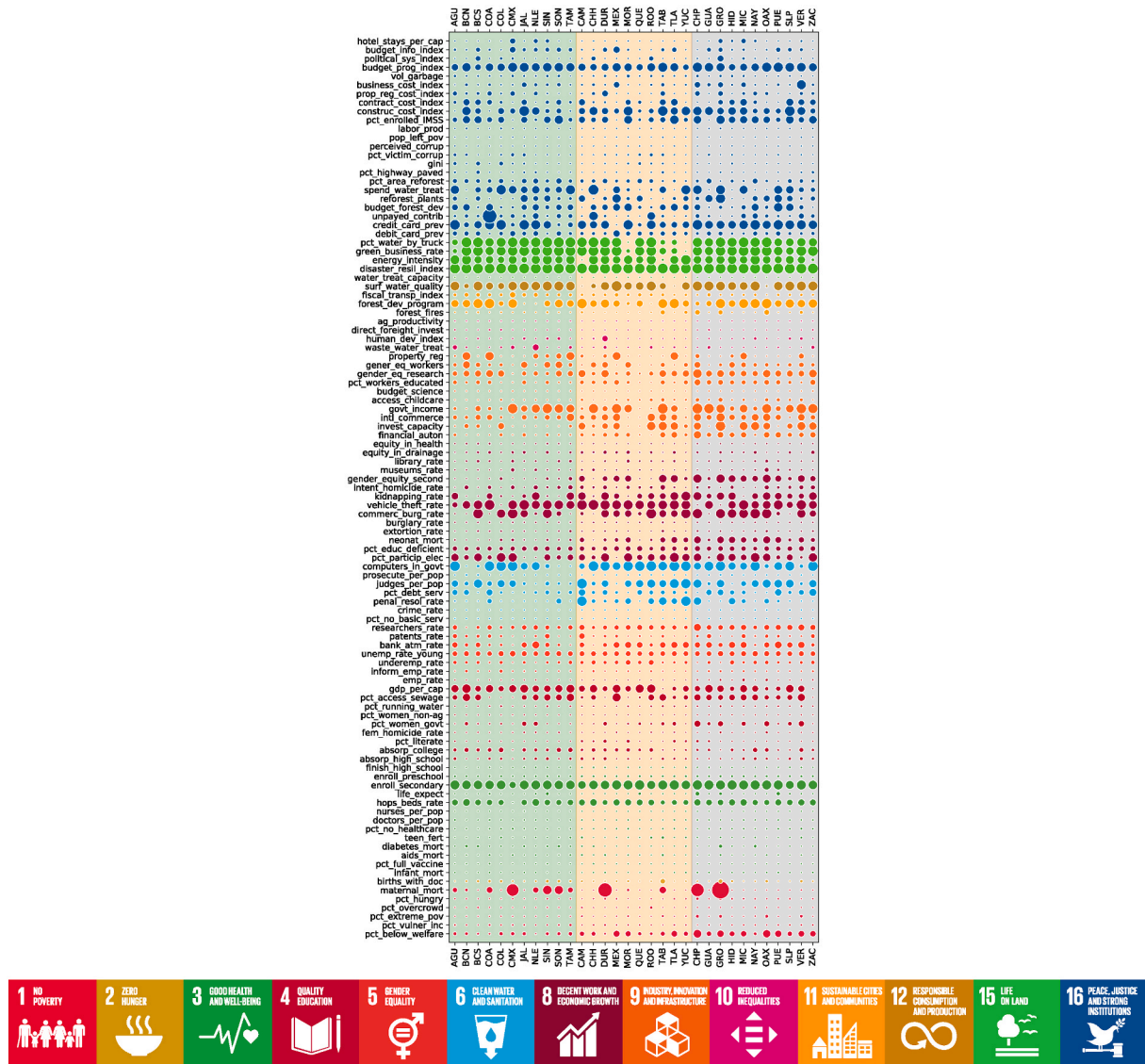

Figure E.1 Development gaps forecasted for 2030 by state and indicator. *Note*: Each dot size is proportional to the forecasted gap size. The background colors correspond to the three clusters. *Source*: Authors' own calculations.

F Constrained optimization of participations

As specified in the Fiscal Coordination Act, the contributions are tied to specific 'themes', while the participations follow a particular formula. With full information, such features could be easily coded into the optimization algorithm. For the purpose of this paper, and to avoid excessive specificities, we provide an example of how to restrict the fitness landscape by considering some of the aforementioned criteria. For this exercise, we assume that contributions remain the same as the empirical ones since they obey specific policy purposes. Thus, the redistribution takes place only in the participations.

The Fiscal Coordination Act defines a weight vector $w_1, \ldots, w_{32}$ where $w_k = 0.6C_{1,k} + 0.3C_{2,k} + 0.1C_{3,k}$ and w_k denotes the weight given to state k. In this formula, the terms $C_{1,k}$, $C_{2,k}$ and $C_{3,k}$ are called the 'incentive coefficients', as they build on information about economic performance (through the state-level GDP for $C_{i,k}$) and local tax-collection performance (in relative changes for $C_{2,k}$ and magnitudes for $C_{3,k}$). These coefficients also consider the state populations, and their full formulations can be found in HCU [35].[30] The exercise consists of introducing a restriction that aligns the optimal solutions to the principle of awarding states with good economic and tax-collection performances. We do this by forcing each proposed solution to preserve the same order as the weight vector $w_1, \ldots, w_{32}$. In other words, if the weights dictate that k should be the most incentivized state, then all the proposed solutions of the optimization algorithm preserves k as the top state in participations received.

Notice, in Figure F1, that the optimal allocation under this constraint is, as expected, different to that obtained in the panel (b) of Fig. 10. In the latter, the algorithm is allowed to test tentative solutions in which participations are moved freely between states. In this new setting the State of

---

[30] With the aim to avoid severe political and economic disruptions, the complete formula is applied only to the additional shareable tax revenues and, hence, the resources preceding those of 2007 are distributed according to the states' shares originally specified for that year. The idea of this formulation is that, as time goes by, the weight of the 2007 distribution will diminish.

Mexico (MEX) is, again, over-assigned. Another salient feature is that Mexico City (CMX) presents, in this case, meaningful transfers. Moreover, the top four states in terms of their empirical participations happen to be over-allocated in the optimal solution. These simulation procedures combine some form of top-down optimal control with the introduction of bottom-up incentives, such as the fostering of local tax recollection.
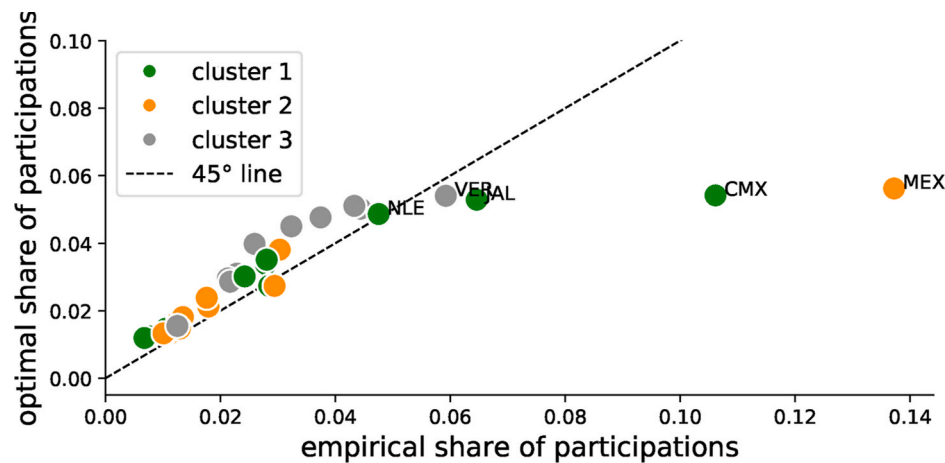

Figure F.1
Constrained optimization of the distribution of participations. *Source*: Authors' own calculations.

References

[1] Forman K, Dougherty S, Blöchliger H. Synthesising good practices in fiscal federalism: key recommendations from 15 Years of country surveys. Paris: OECD Publishing; 2020.

[2] Ivanyna M, Shah A. How close is your government to its people? Worldwide indicators on localization and decentralization. Economics: The Open-Access 2014; 8(2014–3):1.

[3] Bruyninckx H, Happaerts S, Brande K, Brande K, editors. Sustainable development and subnational governments: policy-making and multi-level interactions. Basingstoke: Palgrave Macmillan UK; 2012.

[4] Happaerts S. Sustainable development and subnational governments: going beyond symbolic politics? Environmental Development 2012;4:2–17.

[5] Lucci P, Lynch A. The SDGs at city level: Mumbai's example. Working Paper of the Overseas Development Institute; 2016. p. 432.

[6] Kawakubo S, Murakami S, Ikaga T, Asami Y. Sustainability assessment of cities: SDGs and GHG emissions. Build Res Inf 2018;46(5):528–39.

[7] Patole M. Localization of SDGs through disaggregation of KPIs. Economies 2018;6 (1):15.

[8] Diaz-Cayeros A. Fiscal federalism and redistribution in Mexico. In: Greer S, Trench A, Elliott H, editors. Federalism in good times and bad; 2021. volume forthcoming.

[9] Oates W. Fiscal federalism. New York, N.Y: Harcourt Brace Jovanovich; 1972.

[10] Prud'homme R. The dangers of decentralization. World Bank Res Obs 1995;10(2): 201.

[11] Tanzi V. Fiscal federalism and decentralization: a review of some efficiency and macroeconomic aspects. Washington, D.C: World Bank; 1995.

[12] Martinez-Vazquez J, Lago-Peñas S, Sacchi A. The impact of fiscal decentralization: a survey. J Econ Surv 2017;31(4):1095–129.

[13] Stossberg S, Blöchliger H. Fiscal decentralisation and income inequality: empirical evidence from OECD countries. Jahrb Natl Stat 2017;237(3):225–73.

[14] Dougherty S, Akgun O. Globalisation, decentralisation and inclusive growth. In: Kim J, Dougherty S, editors. *Fiscal Decentralisation and inclusive growth*, pages 49–73. Paris: OECD and Korea Institute of Public Finance; 2018.

[15] Dougherty S, Lorenzoni L, Marino A, Murtin F. The impact of decentralisation on the performance of health care systems: a non-linear relationship. OECD Publishing; 2019.

[16] Lastra-Anadón C, Mukherjee S. Cross-country evidence on the impact of decentralisation and School autonomy on educational performance. OECD Publishing; 2019.

[17] Arechederra F, Urzúa C. La Ley de Coordinación Fiscal en México: Una Crítica Aritmética. Sobre México: Temas de Economía 2017;3(1).

[18] Sánchez M, Ballinez R. Implicaciones de la Ley de Coordinación Fiscal en la Distribución de Recursos Hacia las Entidades Federativas. Technical report. HR Credit Rating Agency; 2020.

[19] Chiguil F. El Sistema Nacional de Coordinación Fiscal (SNCF) y sus Efectos en el Endeudamiento y el Esfuerzo Fiscal del Distrito Federal y de las Entidades del País, 1995-2010. Análisis Económico 2014;29(71):79–111.

[20] Hernández F, Rabling B. Transferences Condicionadas Federales en Países en Desarrollo: el Caso del FISM en México. Estud Económicos 2007;22(2):143–84. 44.

[21] Guerrero O, Castañeda G. How does government expenditure impact sustainable development? Studying the multidimensional link between budgets and development gaps. SSRN Working Paper; 2021.

[22] Guerrero O, Castañeda G. Quantifying the coherence of development policy priorities. Development Policy Review; 2020.

[23] Guerrero O, Castañeda G. Policy priority inference: a computational framework to analyze the allocation of resources for the sustainable development goals2. Data & Policy; 2020.

[24] Castañeda G, Chávez-Juárez F, Guerrero O. How do governments determine policy priorities? Studying development strategies through networked spillovers. J Econ Behav Organ 2018;154:335–61.

[25] Castañeda G, Guerrero O. The resilience of public policies in economic development. Complexity; 2018. 2018.

[26] Castañeda G, Guerrero O. The importance of social and government learning in ex ante policy evaluation. J Pol Model 2019;41(2):273–93.

[27] Guerrero Omar A, Castañeda Gonzalo. Does Expenditure in Public Governance Guarantee Less Corruption? Non-Linearities and Complementarities of the Rule of Law. Economics of Governance 2021;22(2). 139–64.

[28] Ospina-Forero L, Castañeda G, Guerrero O. Estimating networks of sustainable development goals. Information & Management; 2020. p. 103342.

[29] Aragam B, Gu J, Zhou Q. Learning large-scale Bayesian networks with the sparsebn package. J Stat Software 2019;91(1):1–38.

[30] OECD. A territorial approach to the sustainable development goals. OECD Publishing; 2020.

[31] Lafortune G, Fuller G, Moreno J, Schmidt-Traub G, Kroll C. SDG index and dashboard: detailed methodological paper, vol. 1. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN); 2018.

[32] Smits J, Permanyer I. The subnational human development database. Scientific Data 2019;6(1):190038.

[33] de Wolff T, Cuevas A, Tobar F. MOGPTK: the multi-output Gaussian process Toolkit. Neurocomputing; 2020.

[34] Jones B, Baumgartner F, Breunig C, Wlezien C, Soroka S, Foucault M, François A, Green-Pedersen C, Koski C, John P, Mortensen P, Varone F, Walgrave S. A general empirical law of public budgets: a comparative analysis. Am J Polit Sci 2009;53(4): 855–73.

[35] HCU. Ley de coordinación fiscal. 2018.

[36] Giugale M, Webb S. Achievements and challenges of fiscal decentralization : lessons from Mexico. Washington DC: World Bank Group; 2000.

[37] Wiesner E. Fiscal Federalism in Latin America: From Entitlements to Markets. Washington, D.C.: IDB; 2003.

Omar A. Guerrero is Head of Computational Social Science Research at The Alan Turing Institute. He has a Ph.D. in Computational Social Science from George Mason University. Previously, he was a fellow at the Oxford Martin School, the Saïd Business School and the Institute of New Economic Thinking at the University of Oxford. His work spans development economics, political economy, sustainable development, housing economics, firm and labor dynamics, and the improvement of policymaking through computational analysis.

Gonzalo Castañeda is currently a professor-researcher in the Economic Division at Centro de Investigación y Docencia Económica (CIDE), Mexico City. He has a Ph.D. in Economics from Cornell University. He is also a member of the National Research System, level III (i. e., the top rank granted by the National Council of Science and Technology, CONACYT). Prof. Castañeda works on Complex Adaptive System with issues related to economic development and just wrote a new book with the title "The Paradigm of Social Complexity: An Alternative Way of Understanding Societies and their economies."

Lucy Hackett is a PhD student in the Agricultural and Resource Economics department at UC Berkeley, focused on Development and Environmental Economics. Her research interests center on how to best manage resources and the environment for sustainable and equitable growth in the developing world; in particular, in empirically measuring how externalities in urban planning and agricultural policy affect water and air quality, and the consequences of these for human health and wellbeing. Her previous research experience has focused on agent-based modelling and econometric estimation of the effect of policy on birth outcomes in Mexico.

Georgina Trujillo is a PhD student in Population Studies at the Center for Demographic, Urban and Environmental Studies, from El Colegio de México. She was senior analyst in the Simulation Unit at Mexico's National Laboratory for Public Policy from 2017 to 2020. Her main research interests are applied econometrics, agent-based modelling and gender inequality.

Florian Chávez-Juárez is co-founder and managing director of CORESO—Collaborative Research Solutions in Switzerland and associate researcher at the National Laboratory of Public Policy in Mexico City. Previously he was assistant at the National Laboratory of Public Policy (LNPP) at the Centre for Center for Research and Teaching in Economics (CIDE) in Mexico City and head of LNPP's Simulation Unit. His main research interests include health economics, inequality of opportunity, capability approach, agent-based modelling and public policy simulation.