Online Appendix: Policy Priority Inference for Sustainable Development

Omar A. Guerrero and Gonzalo Castañeda

A Data preprocessing

For clarity, let us introduce some notation. Let V denote a matrix where the rows are the units of observation and the columns correspond to periods. Entry $V_{i,t}$ denotes the i^{th} element in period t. Then, if we perform an operation on one of the indices, while holding a specific value for the other, we replace the operated index by the dot symbol $\cdot$. For example, if we want the maximum value across time for the i^{th} value of V, then we write $\max(V_{i,\cdot})$. Similarly, if we want the minimum value of V in a given period t, then we write $\min(V_{\cdot,t})$. Finally, if we omit the second index it means we have a vector.

A.1 Normalization

We normalize the indicators in the range [0,1]. The purpose of normalizing is to make the indicators comparable. Data expressed in percentages, may not need this normalization. However, if a large sample across countries can be collected, then a normalization would be preferred as it would provide more realistic bounds. This is so because the levels of other countries act as benchmarks for how low and how high can indicators be.

In order to normalize the indicators we employ the standard formula

$$\mathcal{I}_{i,t} = \frac{\mathscr{I}_{i,t} - \min(\mathscr{I}_{i,\cdot})}{\max(\mathscr{I}_{i,\cdot}) - \min(\mathscr{I}_{i,\cdot})}, \tag{1}$$

where $\mathscr{I}$ denotes the raw indicator and $\mathcal{I}$ the normalized one. The min and max operators are applied to the entire time series of indicator i across all available countries in the sample. In the case of Mexico, the data have been normalized across a larger sample covering 298 countries and territories for 27 years.

A.2 Imputation

PPI requires the initial and the final values of each indicator. However, should the user want to estimate the spillover network through quantitative methods, then it is desirable to also have observations in-between. While the data collected for this study has comprehensive time-series coverage in the sampling period, there are still some missing observations. To remedy this problem, we generate linear interpolations.

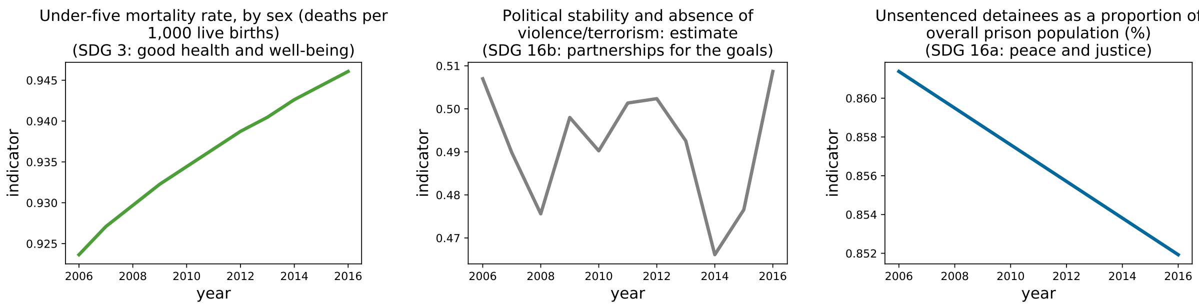
A.3 Reversing

For the purpose of an easier interpretation, it is desirable that the higher values of an indicator denote better outcomes. Since we normalized the indicators in the range [0,1], we can reverse them by using the complement $1 - \mathcal{I}_{i,t}$.

A.4 Target adjustment

The model underlying PPI guarantees convergence to a vector of targets T. It does not matter if those targets are above or bellow the initial values of the indicators. Assuming that the indicators have been pre-processed as suggested above, and that higher values imply better outcomes, convergence to lower values would be inconsistent with the model's logic. This is so because negative dynamics would imply that the government's systematic investment drives the indicators to worse outcomes. Moreover, the spillover network should already account for the negative externalities. Hence, the target of an indicator should always be greater than its initial value.

Figure 1: Development indicators exhibit diverse dynamics



The indicators have been normalized in the interval [0,1] and have been reversed if necessary. Thus, higher values indicate better outcomes.

Figure 1 shows how different development indicators show various dynamics. PPI simplifies the challenge of modeling such diverse patterns through its stochastic growth process. However, this does not fix the problem of having indicators with final values lower than the initial ones. To remedy this issue for retrospective estimations, we propose an adjustment of the retrospective goals using the formula

$$T_i = I_{i,m} + |\min(I_{\cdot,m} - I_{\cdot,1})| + \epsilon, \tag{2}$$

where m is the final period in the sample and $\epsilon > 0$ is a small term close to zero. This calculation needs to be performed for all indicators, even if only one of them exhibits $I_{i,m} < I_{i,1}$

Equation 2 shifts all the final values upwards, guaranteeing $T_i > I_{i,1}$ for every i. Effectively, it assigns the smallest historical gap to the worst-performing indicator, and the largest one to the best-performing one. Thus, the historical gaps capture how much progress was achieved in each indicator during the sampling period.

Besides the argument of model consistency, a second motivation for this adjustment is linked to the growth factors α. Assume there are no network effects, and that the inefficiencies and allocations are the same across all nodes. Then, the only parameter that could explain the variation between the historical gaps is α_i. Since all the empirical indicators arrived to their final values at the same time, it must be the case that the worst-performing indicator has the smallest growth factor and the best one has the largest α_i. What we have done here is mapping the performance of the indicators into the growth factors. This is why α_i can be interpreted as everything else that explains the indicator dynamics but that is not explicitly considered in the model.

B Data

We provide the complete list of development indicators in Table 1.

Table 1: Development indicators

Description	Source	ODS	Instrumental	Reversed
Poverty gap at 5.50 dollars a day (2011 ppp) (%)	World Bank	1	yes	yes
Population in moderate poverty	CONEVAL	1	yes	yes
Population in extreme poverty	CONEVAL	1	yes	yes
Population that is vulnerable due to poor social capital	CONEVAL	1	yes	yes
Population that is vulnerable due to poor income	CONEVAL	1	yes	yes
Lack of health services	CONEVAL	1	yes	yes
Lack of social security	CONEVAL	1	yes	yes
Lack of quality and space in the dwelling	CONEVAL	1	yes	yes
Lack of basic house services	CONEVAL	1	yes	yes
Plant breeds for which sufficient genetic resources are stored (number)	UN	2	yes	no
Proportion of local breeds classified as being at unknown level of risk of extinction (%)	UN	2	yes	yes
Cereal yield (kg per hectare)	World Bank	2	no	no
Food production index (net, per capita)	FAO	2	yes	no
Prevalence of anemia among women of reproductive age (% of women ages 15-49)	World Bank	2	yes	yes
Under-five mortality rate, by sex (deaths per 1,000 live births)	UN	3	yes	yes
Number of new hiv infections per 1,000 uninfected population, by sex and age (per 1,000 uninfected population)	UN	3	no	yes
Tuberculosis incidence (per 100,000 population)	UN	3	yes	yes
Malaria incidence per 1,000 population at risk (per 1,000 population)	UN	3	yes	yes
Mortality rate attributed to cardiovascular disease, cancer, diabetes or chronic respiratory disease (probability)	UN	3	no	yes
Suicide mortality rate, by sex (deaths per 100,000 population)	UN	3	no	yes
Alcohol consumption per capita (aged 15 years and older) within a calendar year (litres of pure alcohol)	UN	3	no	yes
Proportion of the target population with access to 3 doses of diphtheria-tetanus-pertussis (dtp3) (%)	UN	3	yes	no

Table 1: Development indicators

Description	Source	ODS	Instrumental	Reversed
Proportion of the target population with access to measles-containing-vaccine second-dose (mcv2) (%)	UN	3	yes	no
Participation rate in organized learning (one year before the official primary entry age), by sex (%)	UN	4	yes	no
Internet access in schools, 1-7 (best)	World Economic Forum	4	yes	no
Quality of the education system, 1-7 (best)	World Economic Forum	4	yes	no
Quality of primary education, 1-7 (best)	World Economic Forum	4	yes	no
Quality of math and science education, 1-7 (best)	World Economic Forum	4	yes	no
Quality of management schools, 1-7 (best)	World Economic Forum	4	yes	no
Extent of staff training, 1-7 (best)	World Economic Forum	4	no	no
School enrollment, secondary (gross), gender parity index (gpi)	World Bank	4	yes	no
Proportion of seats held by women in national parliaments (% of total number of seats)	UN	5	yes	no
Proportion of women in managerial positions (%)	UN	5	no	no
Water body extent (permanent and maybe permanent) (% of total land area)	UN	6	yes	no
Proportion of population with access to electricity, by urban/rural (%)	UN	7	yes	no
Proportion of population with primary reliance on clean fuels and technology (%)	UN	7	yes	no
Access to clean fuels and technologies for cooking (% of population)	World Bank	7	yes	no
Annual growth rate of real GDP per capita (%)	UN	8	no	no
Number of commercial bank branches per 100,000 adults	UN	8	no	no
Unemployment rate, by sex and age (%)	UN	8	no	yes
Foreign direct investment, net inflows (% of GDP)	World Bank	8	yes	no
Index of economic complexity	Observatory of Economic Complexity	8	no	no
Efficiency of government spending	World Economic Forum	8	yes	no
Burden of government regulation, 1-7 (best)	World Economic Forum	8	yes	no
Burden of customs procedures, 1-7 (best)	World Economic Forum	8	yes	no
Regulation of securities exchanges, 1-7 (best)	World Economic Forum	8	yes	no
Business impact of rules on fdi, 1-7 (best)	World Economic Forum	8	yes	no
Strength of auditing and reporting standards, 1-7 (best)	World Economic Forum	8	yes	no
Protection of minority shareholders' interests, 1-7 (best)	World Economic Forum	8	no	no
Intensity of local competition, 1-7 (best)	World Economic Forum	8	yes	no
Effectiveness of anti-monopoly policy, 1-7 (best)	World Economic Forum	8	yes	no
Extent of market dominance, 1-7 (best)	World Economic Forum	8	yes	no
Efficacy of corporate boards, 1-7 (best)	World Economic Forum	8	no	no
Cooperation in labor-employer relations, 1-7 (best)	World Economic Forum	8	yes	no
Flexibility of wage determination, 1-7 (best)	World Economic Forum	8	yes	no
Pay and productivity, 1-7 (best)	World Economic Forum	8	no	no
Tax revenue (% of GDP)	World Bank	8	yes	no
New business density (new registrations per 1,000 people ages 15-64)	WDI	8	yes	no
Imports as a percentage of GDP	World Economic Forum	8	no	no
Strength of investor protection, 0-10 (best)	World Economic Forum	8	yes	no
Patent applications, residents	World Bank	8	no	no
Contribution of labor quality to GDP growth	The Conference Board	8	no	no
Exports of goods and services (% of GDP)	World Bank	8	no	no
Gdp, ppp (constant 2011 international dollars)	World Bank	8	no	no

Table 1: Development indicators

Description	Source	ODS	Instrumental	Reversed
Wage and salaried workers, total (% of total employment) (modeled ilo estimate)	World Bank	8	no	no
No. days to start a business	World Economic Forum	8	yes	yes
No. procedures to start a business	World Economic Forum	8	yes	yes
Rate of informal employment	INEGI	8	no	yes
Growth of total factor productivity	The Conference Board	8	no	no
Number of fixed internet broadband subscriptions, by speed (number)	UN	9	no	no
Internet users per 100 inhabitants	UN	9	no	no
Manufacturing value added per capita (constant 2010 united states dollars)	UN	9	no	no
Available airline seat km/week, millions	World Economic Forum	9	no	no
Quality of overall infrastructure, 1-7 (best)	World Economic Forum	9	yes	no
Quality of roads, 1-7 (best)	World Economic Forum	9	yes	no
Quality of air transport infrastructure, 1-7 (best)	World Economic Forum	9	yes	no
Quality of electricity supply, 1-7 (best)	World Economic Forum	9	yes	no
Availability of latest technologies, 1-7 (best)	World Economic Forum	9	yes	no
Firm-level technology absorption, 1-7 (best)	World Economic Forum	9	no	no
Fdi and technology transfer, 1-7 (best)	World Economic Forum	9	yes	no
Quality of scientific research institutions, 1-7 (best)	World Economic Forum	9	yes	no
Government procurement of advanced tech products, 1-7 (best)	World Economic Forum	9	yes	no
Soundness of banks, 1-7 (best)	World Economic Forum	9	yes	no
Venture capital availability, 1-7 (best)	World Economic Forum	9	no	no
Financing through local equity market, 1-7 (best)	World Economic Forum	9	yes	no
Availability of research and training services, 1-7 (best)	World Economic Forum	9	yes	no
Company spending on r&d, 1-7 (best)	World Economic Forum	9	no	no
Capacity for innovation, 1-7 (best)	World Economic Forum	9	yes	no
Availability of scientists and engineers, 1-7 (best)	World Economic Forum	9	yes	no
Quality of port infrastructure, 1-7 (best)	World Economic Forum	9	yes	no
Fixed telephone lines/100 pop.	World Economic Forum	9	yes	no
Investment in energy with private participation (current us dollars)	World Bank	9	yes	no
Investment in transport with private participation (current us dollars)	World Bank	9	yes	no
Mobile telephone subscriptions/100 pop.	World Economic Forum	9	no	no
Labour share of GDP, comprising wages and social protection transfers (%)	UN	10	no	no
Ease of access to loans, 1-7 (best)	World Economic Forum	10	yes	no
Income share held by lowest 10%	World Bank	10	no	no
Gini index (world bank estimate)	World Bank	10	no	yes
Pm2.5 air pollution, population exposed to levels exceeding who guideline value (% of total)	World Bank	11	yes	yes
Material footprint per capita, by type of raw material (tonnes)	UN	12	no	yes
Domestic material consumption per capita, by type of raw material (tonnes)	UN	12	no	no
Degree of customer orientation, 1-7 (best)	World Economic Forum	12	no	no
Ethical behavior of firms, 1-7 (best)	World Economic Forum	12	no	no
Adjusted net savings, excluding particulate emission damage (% of gni)	World Bank	12	yes	no
Coal rents (% of GDP)	World Bank	12	no	yes

Table 1: Development indicators

Description	Source	ODS	Instrumental	Reversed
Forest rents (% of GDP)	World Bank	12	no	yes
Mineral rents (% of GDP)	World Bank	12	no	yes
Natural gas rents (% of GDP)	World Bank	12	no	yes
Oil rents (% of GDP)	World Bank	12	no	yes
Total natural resources rents (% of GDP)	World Bank	12	no	yes
Intensity of emissions, meat and cattle	FAO	13	yes	yes
Temperature variation	FAO	13	no	yes
Average proportion of marine key biodiversity areas (kbas) covered by protected areas (%)	UN	14	yes	no
Average proportion of terrestrial key biodiversity areas (kbas) covered by protected areas (%)	UN	15	yes	no
Average proportion of mountain key biodiversity areas (kbas) covered by protected areas (%)	UN	15	yes	no
Red list index	UN	15	yes	no
Unsentenced detainees as a proportion of overall prison population (%)	UN	16a	yes	yes
Business costs of terrorism, 1-7 (best)	World Economic Forum	16a	no	no
Business costs of crime and violence, 1-7 (best)	World Economic Forum	16a	yes	no
Organized crime, 1-7 (best)	World Economic Forum	16a	yes	no
Reliability of police services, 1-7 (best)	World Economic Forum	16a	yes	no
Intentional homicides (per 100,000 people)	World Bank	16a	yes	yes
Public trust in politicians, 1-7 (best)	World Economic Forum	16b	no	no
Favoritism in decisions of government officials, 1-7 (best)	World Economic Forum	16b	yes	no
Transparency of government policymaking, 1-7 (best)	World Economic Forum	16b	yes	no
Property rights, 1-7 (best)	World Economic Forum	16b	yes	no
Intellectual property protection, 1-7 (best)	World Economic Forum	16b	yes	no
Judicial independence, 1-7 (best)	World Economic Forum	16b	yes	no
Government effectiveness: estimate	World Bank	16b	yes	no
Overall level of statistical capacity (scale 0 - 100)	World Bank	16b	yes	no
Legal rights index, 0-10 (best)	World Economic Forum	16b	yes	no
Political stability and absence of violence/terrorism: estimate	World Bank	16b	yes	no
Regulatory quality: estimate	World Bank	16b	yes	no
Corruption perception index	Transparency International	16b	no	no
Voice and accountability: estimate	World Bank	16b	yes	no
Debt service as a proportion of exports of goods and services (%)	UN	17	yes	yes
Prevalence of foreign ownership, 1-7 (best)	World Economic Forum	17	no	no
Prevalence of trade barriers, 1-7 (best)	World Economic Forum	17	yes	no
Gross national savings, % GDP	World Economic Forum	17	no	no
Inflation, annual % change	World Economic Forum	17	no	yes
Travel and tourism direct contribution to GDP percentage share of total GDP	World Travel & Tourism Council	17	yes	no

Table 2: Indicators with missing observations

Indicator	SDG	Missing years
Poverty gap at 5.50 dollars a day	1	2007, 2009, 2011, 2013, 2015
Population in moderate poverty	1	2007, 2009, 2011, 2013, 2015
Population in extreme poverty	1	2007, 2009, 2011, 2013, 2015
Population that is vulnerable due to poor social capital	1	2007, 2009, 2011, 2013, 2015
Population that is vulnerable due to poor income	1	2007, 2009, 2011, 2013, 2015
Lack of health services	1	2007, 2009, 2011, 2013, 2015
Lack of social security	1	2007, 2009, 2011, 2013, 2015
Lack of quality and space in the dwelling	1	2007, 2009, 2011, 2013, 2015
Lack of basic house services	1	2007, 2009, 2011, 2013, 2015
Plant breeds for which sufficient genetic resources are stored	2	2006, 2007, 2008, 2009, 2011, 2013, 2015
Malaria incidence per 1,000 population at risk	3	2006, 2007, 2008, 2009, 2011, 2012, 2013, 2014
Mortality rate attributed to cardiovascular disease, cancer, diabetes or chronic respiratory disease	3	2006, 2007, 2008, 2009, 2011, 2012, 2013, 2014
Suicide mortality rate, by sex	3	2006, 2007, 2008, 2009, 2011, 2012, 2013, 2014
Alcohol consumption per capita	3	2006, 2007, 2008, 2009, 2011, 2012, 2013, 2014
Proportion of women in managerial positions	5	2006, 2007, 2008, 2009, 2010, 2011, 2012
Proportion of population with primary reliance on clean fuels and technology	7	2006, 2007, 2008, 2009, 2011, 2012, 2013, 2014
Annual growth rate of real GDP per capita	8	2008, 2009, 2013
Unemployment rate, by sex and age	8	2006
Tax revenue	8	2006, 2007

Table 2: Indicators with missing observations

Indicator	SDG	Missing years
Investment in energy with private participation	9	2009
Income share held by lowest 10%	10	2007, 2009, 2011, 2013, 2015
Gini index	10	2007, 2009, 2011, 2013, 2015
Pm2.5 air pollution, population exposed to levels exceeding who guideline value	11	2006, 2007, 2008, 2009
Unsentenced detainees as a proportion of overall prison population	16a	2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015

C Model estimation and calibration

This section provides the full details on how to calibrate the model's free parameters and how to perform Monte Carlo simulations to generate inference.

C.1 Estimation of growth factors

The model considers a vector $\alpha_1, \ldots, \alpha_N$ of growth factors as part of the indicators' dynamics. The purpose of estimating each individual growth factor is to homogenize convergence times across indicators. That is, since all empirical indicators reach their final values in the same number of periods, we seek to choose a vector of αs that preserves this property in the synthetic indicators.

Finding the vector of growth factors can prove challenging because of the interdependencies between indicators. For for instance, increasing α_i may affect the convergence time of other indicators through the network because i's 'steps' become larger and so do the spillovers. Thus, simultaneously estimating all αs is not a trivial problem, and we have found that many non-linear optimization methods fail. For this reason, we have devised a heuristic to solve this problem.

Our estimation method computes the marginal effect of each growth factor independently until an error term is minimized. In order to think about the error, let us assume that we want all indicators to converge simultaneously after $\mathcal{T}$ periods. Then, the objective is to minimize the average deviation from $\mathcal{T}$ across all indicators and simulations.

First, let us determine an arbitrary vector of growth factors. Using these factors, we perform one simulation run until all indicators converge; then, we obtain a vector with the number of periods that it took each indicator to converge. By repeating this step m times, we obtain m convergence time vectors, which allows computing the average convergence time V_i for each indicator. Next, we compute the convergence error

$$e_v = \frac{1}{N} \sum_{i}^{N} |\mathcal{T} - V_i|. \tag{3}$$

Next, we want to identify those indicators whose convergence error $|\mathcal{T} - V_i|$ is greater than e_v. For one of these indicators, say i, we we vary α_i marginally. Then, we perform m simulations and compute $|\mathcal{T}-V_i|$. We repeat these two steps for i, covering the range $(0, 1)$ for α_i. Then, we choose the growth factor that minimizes i's the difference between convergence error and e_v. This greedy search is performed for each indicator with a convergence error $|\mathcal{T} - V_i| > e_v$, until we obtain a new vector of convergence factors. With this new vector, we re-estimate e_v and repeat all the previous steps. The procedure stops when an error threshold has been reached for e_v.

While our heuristic assumes a *ceteribus paribus* condition in every greedy search, it is effective in finding the growth factors. Since V_i consists of average convergence times, the error is sensitive to size of m. That is, larger ms decrease the variance of convergence times. Consequently, the mapping from α_i and V_i during the greedy search becomes more accurate with more simulations. This of course, comes with a computational burden. Therefore, model estimation greatly benefits from parallel processing. Algorithm 1 shows the pseudocode of the estimation procedure.

Algorithm 1: Estimation pseudocode

Input: $\mathcal{T}$, $\alpha_1, \ldots, \alpha_N$, initial I, T, $\mathbb{A}$, φ, τ, m

```
while e_v > p do
    foreach i such that |T − V_i| > e_v do
        set α_i = argmin|T − V_i(α_i)|;
    update convergence error e_v;
```

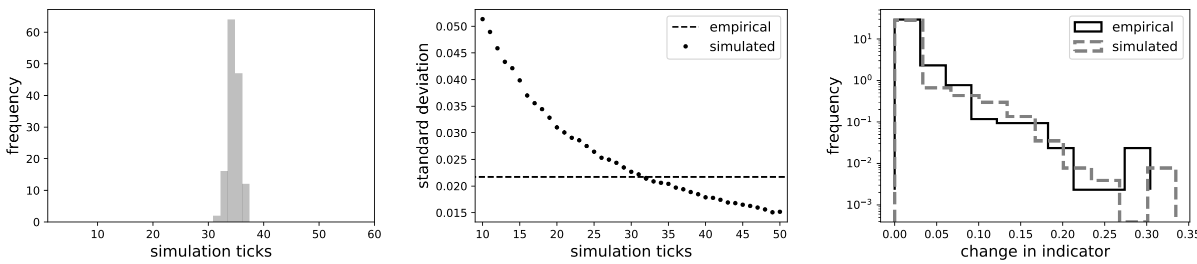
C.2 Calibration of simulation periods

Our estimation procedure assumes a number of periods $\mathcal{T}$ under which the model should converge. In order to calibrate $\mathcal{T}$, we aim at the third data feature discussed in section 4 of the main text: matching the total volatility of the indicators.

First, it is necessary to adjust the empirical data so that its volatility is comparable to the synthetic one. This is so because, while the empirical data may exhibit upward and downward dynamics, PPI's model only generates growth dynamics. Therefore, the adjustment consists of computing one-period changes in all the empirical indicators and, then, turning any negative change into zero. Now, we can calculate the standard deviation of the adjusted data. To calibrate $\mathcal{T}$, we need to find a number of periods under which the estimated growth factors yield indicator dynamics with similar volatility to the empirical one. This means that the entire estimation procedure needs to be performed every time a different $\mathcal{T}$ is chosen.

Figure 2 shows the results of the estimation and calibration procedures. The left panel presents a histogram of average convergence times across indicators once the model has been estimated and calibrated. Clearly, the estimated growth factors for $\mathcal{T} = 32$ generate a small divergence from the target number of periods. Here, the average error e_v is less than one. The middle panel shows the volatility of the simulated indicators obtained for different levels of $\mathcal{T}$ (each one with its estimated growth factors). For this study, $\mathcal{T} = 32$ yields the best match between the empirical and the simulated volatility. In the right panel we can see the histogram of the changes in the indicators (one for the empirical data and one for the simulated). Once estimated and calibrated, the model generates a similar distribution to the empirical one.

Figure 2: Model estimation and calibration



Model estimation for a calibration of $\mathcal{T} = 32$. Left: histogram of average convergence times across indicators. Middle: indicator volatility as a function of $\mathcal{T}$. Right: empirical and fitted distributions of the changes in the indicators.

C.3 Computational efficiency

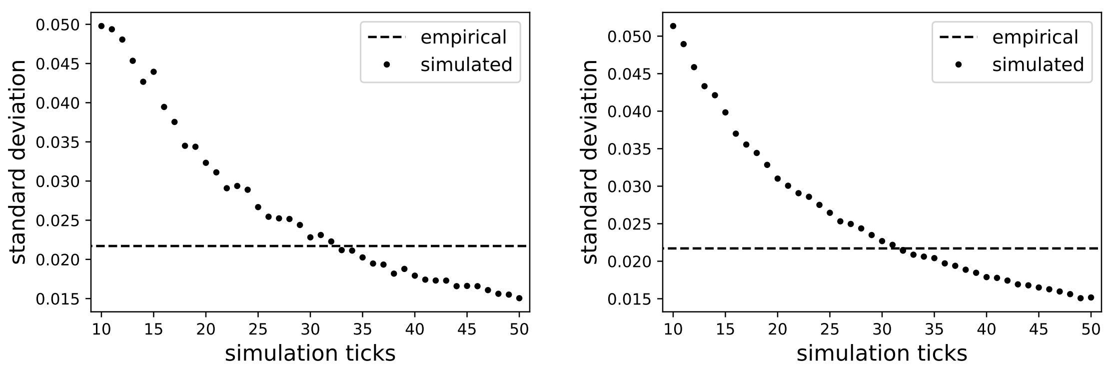
Combined, the estimation and calibration procedures can be computationally expensive. As we have mentioned earlier, one way to reduce the computational burden is to run these processes in parallel. However, through our understanding about the model, we can provide further advice on how to manage the computational burden of estimating the growth factors.

We have found that the volatility of the simulated indicators is not too sensitive to the precision with which the growth factors are estimated, but rather to assumed the number of convergence periods $\mathcal{T}$. In other words, as an initial step to find an optimal $\mathcal{T}$, one may relax the convergence time error in equation 3 in order to produce a mapping like the one in the middle panel of figure 2. By relaxing the error threshold, we do not need to run numerous simulations during the greedy search of algorithm 1, significantly reducing the computational burden. Once the mapping between $\mathcal{T}$ and the indicators' volatility has been produced, we can focus in the optimal $\mathcal{T}^*$. Then, we can re-estimate the growth factors with more simulations and a more conservative error. As a verification step, one may repeat this last re-estimation with $\mathcal{T}^* - 1$ and $\mathcal{T}^* + 1$ to check that the best volatility is still given by

$\mathcal{T}^*$. We have found that this to be a very effective strategy to increase the computational efficiency of our estimation method.

Figure 3 shows the result of the calibration procedure under different computational burdens. Clearly, while there is some sensitivity to the number of simulations ran in the greedy search of the estimation method, the overall mapping of $\mathcal{T}$ into volatility is robust, yielding $\mathcal{T}^* = 32$ for the implementation with fewer simulations.

Figure 3: Reduction in computational burden



Left: calibration procedure running 10 simulations in the greedy search. Right: calibration procedure running 100 simulations in the greedy search.

C.4 Monte Carlo simulation

Once the model has been estimated and calibrated, PPI can be used to produce a number of inferences via Monte Carlo simulation. It is important to run independent realizations of the model because the uncertain environment under which the agents learn may lead to decision paths that are specific to a particular simulation. The idea behind the Monte Carlo approach is to generate many realizations of the world, and to compute the expected values of the variables of interest. Besides the model parameters, it is also necessary to initialize the endogenous variables in a random fashion (e.g., P, H, X, etc.). Thus, in order to account for model uncertainty,on could construct confidence intervals from these distributions. The more simulations performed, the narrower those intervals become. In other words, inferences become more robust with a large number of simulations.

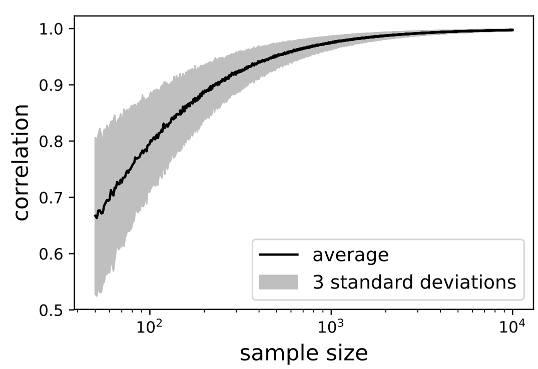
To demonstrate this point, let us concentrate on the main endogenous variable: the allocation profile P. Recall that the inferred allocation profile is built by, first, computing the inter-temporal sum of allocations in each indicator and, then, calculating the average of this quantity across simulations for each indicator. More formally, the expected allocation profile is

$$\mathscr{P}_i = \frac{1}{M} \sum_m \sum_t P_{i,t,m}, \tag{4}$$

where m is the index of the m^{th} out of M Monte Carlo simulations.

In order to assess the robustness, we measure the similarity between two expected allocation profiles: one obtained from with M and another calculated from $M + 1$ simulations. That is, for two samples of sizes M and $M + 1$, we compute multiple pairs of expected allocation profiles and calculate their Spearman correlation. If sample size improves the estimation, then the variation of the Spearman correlation should decrease as we increase M. Figure 4 confirms this. After performing 1000 simulations, the estimated policy priorities are robust.

Figure 4: Robustness of allocation profiles across different sample sizes



D Network statistics

Table 3 provides the summary statistics of the spillover network estimated for this study.

Table 3: Summary statistics of SDG network

SDG or Pillar	In(Out)-Degree Synergies	In(Out)-Degree Trade-Offs	In(Out)-Strength Synergies	In(Out)-Strength Trade-Offs
1	2 (0)	3 (1)	1.46 (0)	1.94 (0.52)
2	14 (10)	17 (14)	11.26 (7.92)	13.93 (11.17)
3	26 (10)	23 (15)	19.17 (7.91)	17.68 (10.78)
4	19 (26)	18 (12)	14.93 (18.76)	14.03 (8.47)
5	9 (2)	8 (3)	6.83 (1.64)	5.74 (2.32)
6	5 (20)	5 (17)	3.76 (16.27)	3.86 (14.28)
7	11 (10)	8 (3)	7.74 (6.75)	5.98 (2.15)
8	61 (47)	31 (25)	47.0 (35.86)	21.39 (18.51)
9	63 (72)	36 (29)	48.77 (55.94)	26.01 (21.62)
10	12 (6)	7 (13)	9.55 (4.57)	4.77 (9.67)
11	3 (0)	2 (0)	2.55 (0)	1.26 (0)
12	8 (17)	8 (18)	5.8 (12.34)	5.42 (12.74)
13	1 (0)	0 (2)	0.73 (0)	0 (1.22)
14	3 (1)	0 (0)	1.68 (0.87)	0 (0)
15	3 (10)	6 (9)	2.22 (7.07)	4.6 (6.95)
16	6 (8)	6 (9)	4.31 (6.12)	4.63 (6.68)
17	3 (10)	2 (8)	2.33 (7.48)	1.24 (5.09)
18	22 (22)	11 (13)	16.5 (17.08)	8.77 (9.09)
social	40 (18)	37 (31)	31.64 (14.13)	27.64 (23.68)
environmental	23 (41)	19 (31)	16.12 (30.96)	14.45 (24.61)
economic	135 (146)	77 (80)	103.9 (111.62)	54.06 (57.95)
human capital	45 (36)	41 (27)	34.1 (26.66)	31.71 (19.25)
institutional	28 (30)	17 (22)	20.81 (23.2)	13.41 (15.77)

The statistics have been separated into synergies and trade-offs. In(out)-degree is the number of incoming(outgoing) connections to(from) a node. The in(out)-strength is the sum of the weights of all incoming(outgoing) connections to(from) a node.