Online Annex 1.1 Datasets for Public Spending[1]

This Online Annex describes the underlying data used in the main text. It summarizes data sources, coverage, and methodology for (i) public spending, (ii) outcome variables for public spending efficiency estimates, and (iii) drivers of public spending efficiency, composition, and rigidity.

Public Spending

The database combines different sources to construct a new series for the level of public spending, both at the aggregate level and for public spending on investment, education, health, and R&D.[2] To ensure consistency, all public spending indicators are calculated first as a share of gross domestic product (GDP) in nominal terms and then converted to real per capita units (for the efficiency gaps analysis—see Online Annex 1.2) using GDP deflators and PPP exchange rates and to percent of total expenditure (for the growth regressions analysis—see Online Annex 1.5). Variable descriptions, sources, coverage, and time periods for these indicators are summarized in Online Annex Table 1.1.1; steps for extraction and calculations are summarized below.

Total public expenditure

Total public expenditure is measured as general government expenditure as a share of GDP and is largely based on the IMF's Public Finances in Modern History (PFMH) database, adjusted for structural breaks.[3] The adjusted series was extended to 2024 using the October 2024 World Economic Outlook (WEO). Any "jumps" introduced by combining data sources were addressed by applying the difference between the two series to the following or preceding years.

Public investment

Public investment is measured as gross fixed capital formation of the general government and is taken primarily from the IMF's Investment and Capital Stock Dataset (ICSD) (IMF 2021). ICSD data were carried backward and forward using the comparable series from the following sources (by order of priority): WEO, World Development Indicators (WDI), Government Finance Statistics (GFS), and Barro-Lee (1993).[4] The series are adjusted for structural breaks, assuming no year-to-year change. [5]

Public education spending

Public education spending, as defined by the Classification of Functions of Government (COFOG), includes "expenditure on services provided to individual pupils and students and expenditure on services provided on a collective basis (IMF 2014)." The series on public education spending as a share of GDP combines data from the following sources: Barro-Lee (1993), Gethin (2024), GFS, Our World in Data (OWID), the United Nations Educational, Scientific and Cultural Organization (UNESCO), and WDI,

---

[1] Prepared by Chloe Hyungsun Cho.

[2] Public spending indicators aim to capture spending across all levels of government by using the consolidated data for the general government sector as the benchmark series where available.

[3] Any changes in coverage (from central to general government) were addressed, assuming no year-to-year change.

[4] All raw data was visually assessed, and some data points (e.g., suspiciously large jumps or negative values that are not visible in other sources) were treated as missing.

[5] Structural breaks were identified as suspicious cases of large jumps (i.e., three standard deviations) combined with disagreement between the ICSD and WEO (public) jump sizes and between ICSD (public) and Penn World Table (aggregate) jump sizes.

which are in line with the COFOG definition but may also include self-reported data. The series from UNESCO was adjusted for structural breaks from the change in raw data sources.[6] Gethin (2024) was used as the benchmark series for most countries for the period of 1980–2022.[7] The series was then (i) carried forward using the GFS, UNESCO, and WDI; and (ii) carried backward using the UNESCO, recent WDI, and vintage data from 2005 WDI. OWID and Barro-Lee (1993) were used as additional sources for historical data.

Public health spending

Public health spending is defined as government expenditures on both individual health services and collective services, which cover policy, regulation, and research (IMF 2014). The public health spending series combines data from the following sources: Gethin (2024), GFS, OWID, vintage WDI, and the World Health Organization (WHO). Gethin (2024)[8] was taken as the benchmark series for most countries but replaced with WHO or GFS for select countries. OWID and vintage data from 2005 WDI were used as additional sources for historical data.

Public R&D spending

Public spending on R&D is defined as the gross domestic expenditure on R&D (GERD) financed by the government. Two data sources were used: the Organization for Economic Co-operation and Development (OECD) and UNESCO, with the OECD data taking priority for the OECD countries, given better time coverage and availability of GDP share data.

---

[6] The breaks were identified based on the metadata provided by UNESCO and were adjusted by assuming no year-to-year change.

[7] As Gethin uses general government education expenditure from Eurostat, OECD, CEPAL, IMF, and UNESCO, it was already in line with many other sources of priority. See Gethin (2024) for the methodology.

[8] Gethin uses general government health expenditure from Eurostat, the OECD, CEPAL, the IMF, the WHO, and SPEED, by order of priority, adding historical series from Murray, Govindaraj, and Musgrove (1994). See Gethin (2024).

Online Annex Table 1.1.1 Public Spending Variables

Indicator	Variable description	Source	No. of countries	Period
Total expenditure				
totexp_gdp	Total government expenditure, percent of GDP	Calculated	194	1800-2024
pfmh_exp_adj	Total government expenditure, percent of GDP	PFMH[1]	151	1800-2023
*ggx_gdp_*	General government total expenditure, percent of fiscal year GDP	WEO	194	1956-2029
Public investment				
pubinv_gdp	Public investment, percent of GDP	Calculated	184	1960-2024
pubinv_totexp	Public investment, percent of expenditure	Calculated	183	1960-2024
pubinv_ppp_pc	Public investment, 2021 PPP$, per capita	Calculated	181	1960-2024
igov_gdp	General government investment, percent of GDP	ICSD[2]	174	1960-2019
weo_nfig_gdp	Public gross fixed capital formation, percent of GDP	WEO[2]	160	1963-2029
wdi_gfcf_pub	Public sector GFCF, calculated as the difference between total and private sector GFCF, percent of GDP	WDI	104	1960-2024
gfs_pubinv	Public investment, calculated as the sum of consumption of fixed capital and investment in nonfinancial assets, percent of GDP	GFS	59	1990-2023
bl_invpub	Ratio of public domestic investment to GDP, nominal	Barro-Lee[3]	109	1960-1984
Public education spending				
education_gdp	Government expenditure on education, percent of GDP	Calculated	194	1870-2024
education_totexp	Government expenditure on education, percent of expenditure	Calculated	194	1870-2024
education_ppp_pc	Government expenditure on education, 2021 PPP$, per capita	Calculated	192	1950-2024
un_ed_gov_adj	Government expenditure on education, percent of GDP, adjusted for structural breaks from data source changes	UNESCO	191	1970-2024
gfs_ed_adj	General government expenditure on education, *GF09,* percent of GDP	GFS	82	1972-2023
wdi_ed_gdp	Government education spending, percent of GDP	WDI	191	1970-2023
wdi05_ed_gdp	Public spending on education, total, percent of GDP	2005 WDI	156	1998-2002
owid_ed	Public spending on education, percent of GDP	OWID[4]	191	1870-2023
bl_ed	Ratio of total government expenditure on education to GDP, nominal	Barro-Lee	114	1960-1984
gethin_ed	Government education expenditure, percent of GDP	Gethin	175	1980-2023
Public health spending				
health_gdp	Government expenditure on health, percent of GDP	Calculated	193	1880-2023
health_totexp	Government expenditure on health, percent of expenditure	Calculated	193	1880-2023
health_ppp_pc	Government expenditure on health, 2021 PPP$, per capita	Calculated	191	1950-2023
who_gghed_gdp	Domestic general government health expenditure, percent of GDP	WHO	189	2000-2023
gfs_health	General government expenditure on health, *GF07*, percent of GDP	GFS	82	1972-2023
wdi05_health_gdp	Public health expenditure, percent of GDP	Vintage WDI	180	1998-2002
owid_health	Public health expenditure as a share of GDP	OWID	51	1880-2023
gethin_health	Government health expenditure, percent of GDP	Gethin	175	1980-2023
Public R&D spending				
research_gdp	Government financed expenditure on R&D, percent of GDP	Calculated	131	1981-2024
research _totexp	Government financed expenditure on R&D, percent of expenditure	Calculated	131	1981-2024
research _ppp_pc	Government financed expenditure on R&D, 2021 PPP$, per capita	Calculated	130	1981-2024
oecd_gerd_gov	GERD financed by government, percent of GDP	OECD	47	1981-2024
un_gerd_gov	GERD financed by government, calculated by applying the government financing share to the total GERD, percent of GDP	UNESCO	132	1996-2019
Other categories of public spending				
gfs_def_ord	General government expenditure on defense and public order, calculated as the sum of *GF02* and *GF03*, percent of GDP	GFS	80	1972-2023
gfs_social	General government expenditure on social protection (including pension), *GF10*, percent of GDP	GFS	82	1972-2023
gfs_pension	General government expenditure on pension, calculated as the sum of *GF1002* (old age) and *GF1003* (survivors), percent of GDP	GFS	65	1990-2023
gfs_research	General government expenditure on R&D in select sectors, calculated as the sum of *GF0408* (economic affairs), *GF0705* (health), *GF0907* (education), and *GF1008* (social protection), percent of GDP	GFS	65	1990-2023

Source: IMF staff calculations.
Note: Missing values in the original series were linearly interpolated.
1 Adjusted for breaks in coverage (i.e., from central to general government).
2 Calculated using the nominal series of each indicator and GDP in the respective database.
3 Five-year average figures were carried across the respective period (e.g., 1960-64 average filled in for all five years).

Outcome Variables for Public Spending Efficiency Estimates

The database captures output or outcome indicators in the areas of interest to evaluate the efficiency of different public spending categories. Where applicable, missing values between any two observed data points for a given country-year series were linearly interpolated. Outcome variables were selected based on previous studies, as well as overall data availability and quality (see Online Annex 1.1.2).

For public investment, most of the data are taken from the WDI, except the total road length used to calculate road density which comes from the International Road Federation (IRF)'s World Road Statistics (WRS). Measures of physical output were complemented by the Global Competitiveness Index (GCI) capturing the quality of overall infrastructure from the World Economic Forum (WEF), downloaded from WDI. For public education spending, the database considers available sources such as Barro-Lee (1993), OECD, UNESCO, WDI, and the World Bank's Education Statistics (EdStats), combining them where possible. For public health spending, it includes measures of healthcare capacity (e.g., the number of physicians per capita) and health outcomes (e.g., life expectancy, infant mortality rates) from WDI and WHO. For public R&D spending, the database captures measures of productivity in research activities using data from Scimago, WDI, and the World Intellectual Property Organization (WIPO).

The list of outcome variables considered for the estimation of public efficiency gaps—with descriptions, sources, coverage, and time periods—is provided in Online Annex Table 1.1.2.

Online Annex Table 1.1.2 Outcome Variables for Public Spending Efficiency Estimates[9]

Indicator	Variable description	Source	No. of countries	Period
Public investment				
elec_consumption_wdi	*Energy*: Electric power consumption (kWh per capita)	WDI	175	1960-2023
elec_access_wdi	*Energy*: Access to electricity, percent of population	WDI	193	1990-2023
tran_rail_wdi_pc	*Transport*: Rail density, calculated as rail lines (total route, km) per 1,000 people	WDI	114	1995-2021
road_density	*Transport*: Road density, calculated as the ratio of total road network (all road types, km) to total land area (km²)	IRF-WRS (road) WDI (land area)	187	1996-2022
paved_ratio_int	*Transport:* Paved Network Ratio (% paved of total roads)	IRF-WRS	181	1996-2022
tran_air_passengers_pc	*Transport*: Air transport, passengers carried per capita	WDI	180	1970-2021
tran_port_traffic_pc	*Transport*: Seaport activity, calculated as container port traffic, TEU (twenty-foot equivalent units) per capita	WDI	153	2000-2022
comm_phone_wdi_int	*Communications:* Fixed telephone subscriptions (per 100 people)	WDI	193	1960-2023
comm_cell_wdi_int	*Communications:* Mobile cellular subscriptions (per 100 people)	WDI	193	1960-2023
sanitation_wdi	*Water and sanitation:* People using at least basic sanitation services, percent of population	WDI	192	2000-2022
water_drinking_wdi	*Water and sanitation:* People using at least basic drinking water services, percent of population	WDI	192	2000-2022
infra_quality	Quality of overall infrastructure (1=worst to 7=best)	WEF	151	2007-2017
Public education				
teacher_ps_pri	Teacher-pupil ratio, primary, calculated as the number of teachers per 1,000 students or the inverse of pupil-teacher ratio	WDI	190	1970-2023
enrollment_pri	Net enrollment rate, primary, percent of relevant age group	UNESCO, WDI & OECD	192	1970-2024
completion_pri	Completion rate, primary, percent of relevant age group	UNESCO & WDI	192	1970-2023
avg_yos	Average years of schooling, population 25+ years	UNESCO, EdStats & Barro-Lee[1]	187	1870-2023
ed_cr_mod_usec_un	Completion rate, upper secondary education, both sexes, percent of relevant age group (modelled data)	UNESCO	163	1990-2024
ed_lr_youth_wdi_int	Literacy rate, percent of population 15-24 years	WDI, EdStats & UNESCO	162	1970-2023
ed_quality	Quality of the education system, 1-7 (best)	WEF	151	2007-2017
Public health				
health_bed_wdi	Hospital beds (per 1,000 people)	WDI	190	1960-2021
health_doc_wdi	Physicians (per 1,000 people)	WDI	192	1960-2022
health_le_wdi	Life expectancy at birth, total (years)	WDI	193	1960-2023
health_DPT_wdi	DPT immunization rate, percent of children ages 12-23 months	WDI	189	1980-2023
health_sr_inf_who	Infant survival rate, calculated as the inverse of infant mortality rate per 1,000 live births (*health_mr_inf_who)*	WHO	190	1932-2023
health_gci	Global Competitiveness Index (GCI) subpillar score on health	WEF	151	2007-2017
health_hale_who	Healthy life expectancy at birth (years), both sexes	WHO	182	2000-2021
health_uhc_who	Uiversal Health Coverage Service Coverage Index (SDG 3.8.1)	WHO	188	2000-2021
Public R&D				
ip_article_pc	Scientific and technical journal articles (per million people)	WDI	189	1996-2022
wipo_patent	Resident patent applications (per million people)	WIPO	145	1980-2023
researcher_pc	Researchers in R&D (per million people)	WDI	136	1996-2023
citations_pc	Publication citations, calculated as the number of citations to all documents published to date (per million people)	Scimago	191	1996-2024

Source: IMF staff calculations.
Note: For indicators that combined data from different sources original sources are listed in the order of priority.
1 Barro-Lee (1993) data was used for a limited sample, as efficiency gaps were calculated for the period 1980-2024.

[9] These indicators may not be direct measures of productivity or may reflect contributions from the private sector or external actors (e.g., donors). Nevertheless, they capture the overall performance and thus serve as a valid benchmark for assessment.

Determinants of Public Spending and Control Variables for Growth Regressions

The list of other variables captured in the database and their descriptions, sources, coverage, and time periods are summarized in Online Annex Table 1.1.3. Where appropriate, variables were transformed using time differences or into logarithmic forms.

Online Annex Table 1.1.3 Drivers of Public Spending and Control Variables

Indicator	Variable description	Source	No. of countries	Period
Demographic				
dependency_old	Age dependency ratio, old, percent of working-age population	WDI	193	1960-2024
wdi_birth_rate	Birth rate, crude (per 1,000 people)	WDI	193	1960-2023
wdi_population	Population, total	WDI	193	1960-2024
wdi_pop1564	Population ages 15-64 (% of total population)	WDI	193	1960-2024
weo_population	Population, persons, millions	WEO	194	1950-2029
Macroeconomic				
figb	Long-term bond yield (Percent, Units)	WEO	59	1950-2029
figb_r	Real long-term bond yield (Percent, Units)	WEO	59	1951-2029
ggxwdg_gdp	General government gross debt, percent of fiscal year GDP	WEO	192	1960-2029
ngap_r	Output gap (Percent, Units)	WEO	106	1950-2029
ngdp_r_ppp	GDP, constant prices, PPP 2021 international dollars	WEO	192	1950-2029
ngdp_r_ppp_pc	GDP, constant prices, PPP 2021 international dollars, per capita	WEO	192	1950-2029
pcpi	Consumer Prices, period average (Index, Units)	WEO	194	1950-2029
proj_ngdppch_1y	Percent change in GDP, current prices, national currency, projection for the year ahead	WEO	194[1]	1991-2029
gini_wiid	Gini index[2] (0=perfect equality to 100=perfect inequality)	WIID	194	1950-2022
emp	Number of persons engaged, millions	PWT	175	1950-2019
rgdpna	Real GDP at constant 2017 national prices, million 2017 USD	PWT	175	1950-2019
rkna	Capital services at constant 2017 national prices (2017=1)	PWT	134	1954-2019
rtfpna	TFP at constant national prices (2017=1)	PWT	118	1954-2019
al_x5	Change in the level of liberalization in telecommunication and electricity sectors, calculated as the first difference of product market indicator (0=highest repression to 1=full liberalization)	Alesina et al. (2024)	90	1974-2014
al_x6	Change in the level of employment protection legislation, calculated as the first difference of labor market liberalization indicator (0=more regulation to 1=no regulation)	Alesina et al. (2024)	90	1974-2014
banking_crisis	Dummy for the starting date of systemic banking crisis (0/1)	Laeven & Valencia (2018)	194	1971-2017
currency_crisis	Dummy for the starting date of currency crisis (0/1)	Laeven & Valencia (2018)	194	1971-2017
debt_crisis	Dummy for the starting date of sovereign debt crisis (0/1)	Laeven & Valencia (2018)	194	1971-2017
debt_restructuring	Dummy for the starting date of sovereign debt restructuring (0/1)	Laeven & Valencia (2018)	194	1971-2017
Policy and governance				
Fiscal policy and public financial management				
rd	Revenue decentralization, calculated as the ratio of own revenue to general government revenue	FD Database	82	1972-2020
sd	Spending decentralization, calculated as the ratio of own spending to general government spending	FD Database	69	1990-2020
transferdep1	Transfer dependency method 1, calculated as the ratio of net transfers to own spending	FD Database	69	1990-2020
transferdep2	Transfer dependency method 2, calculated as the ratio of net transfers to own revenue	FD Database	68	1990-2020
vfi	Vertical Fiscal Imbalance (1- ratio of own revenue to own spending)	FD Database	69	1990-2020
council_dummy	Dummy for the existence of a fiscal council (0/1)	FRFC Database	121	1985-2024
rule_BBR	Dummy for the existence of a budget balance rule (0/1)	FRFC Database	121	1985-2024
rule_DR	Dummy for the existence of a debt rule (0/1)	FRFC Database	121	1985-2024
pima_allo_des	Institutional design of public investment allocation, calculated as the average score for five institutions (1=low, 2=medium, 3=high)	PIMA Database	91	2015-2024
pima_allo_eff	Effectiveness of public investment allocation, calculated as the average score for five institutions (1=low, 2=medium, 3=high)	PIMA Database	91	2015-2024
pima_impl_des	Institutional design of public investment implementation, calculated as the average score for five institutions (1=low, 2=medium, 3=high)	PIMA Database	91	2015-2024
pima_impl_eff	Effectiveness of public investment implementation, calculated as the average score for five institutions (1=low, 2=medium, 3=high)	PIMA Database	91	2015-2024
pima_plan_eff	Institutional design of public investment planning, calculated as the average score for five institutions (1=low, 2=medium, 3=high)	PIMA Database	91	2015-2024
pima_plan_eff	Effectiveness of public investment planning, calculated as the average score for five institutions (1=low, 2=medium, 3=high)	PIMA Database	91	2015-2024

Overall policy and governance				
icrg_qog	Quality of government indicator from the International Country Risk Guide (ICRG); mean value of scores on corruption, law and order, and bureaucracy quality (0=lowest to 1=highest)	QoG	138	1984-2022
functioning_govt	Functioning of government rating from Freedom House; the extent to which government policies are set by elected representatives, free from corruption, open and accountable (0=worst to 12=best)	QoG	189	2005-2021
civil_liberty	Civil liberties rating from Freedom House; freedom of expression and belief, associational and organizational rights, rule of law, and personal autonomy and individual rights (1=most to 7=least free)	QoG	189	1972-2021
rule_of_law	Rule of law rating from Freedom House; the level of judicial independence, prevalence of rule of law, civilian control of police, protection from political violence, absence of conflict, and equal treatment under the law (0=worst to 16=best)	QoG	189	2005-2021
fragility	State fragility index from the Center of Systemic Peace, sum of effectiveness and legitimacy scores in security, political, economic and social dimensions (0=no fragility to 25=extreme fragility)	QoG	164	1995-2018
EGDI	E-Government Development Index, calculated as the average of normalized Online Service Index, Telecommunication Infrastructure Index, and Human Capital Index	United Nations	188	2003-2024
cci	Control of corruption index; perceptions of the extent to which public power is exercised for private gain, normalized scores in standard deviations[3] (i.e., ranging from approximately -2.5 to 2.5, higher scores indicating better control or less perceived corruption)	WGI	193	1996-2023
Political				
dpi_execrlc	Chief executive party orientation (0=no information, 1=right, 2=center, 3=left, n/a=no executive)	DPI	173	1975-2020
dpi_maj	Margin of majority (percent of seats held by the government)	DPI	173	1975-2020
dpi_yrcurnt	Number of years left in current term of chief executive	DPI	172	1975-2020

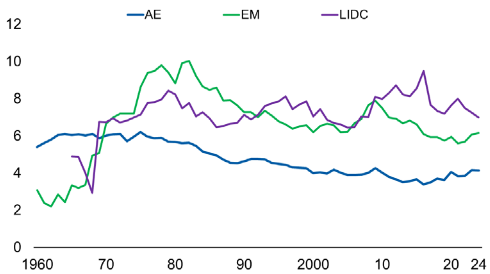
Source: IMF staff calculations.
Note: PWT=Penn World Table; TFP=Total Factor Productivity; DPI=Database of Political Institutions; PIMA=Public Investment Management Assessment; WGI=Worldwide Governance Indicators; QoG=Quality of Government; FD=Fiscal Decentralization; FRFC=Fiscal Rules and Fiscal Council.
1 Vintage forecast data was collected from previous versions of WEO which cover different numbers of countries and years (with smaller samples for older forecasts). The table reflects the overall coverage information.
2 The Gini index is a statistical measure of distribution that represents the income inequality within a population. It is derived from the Lorenz curve, using the area between the curve and a hypothetical line of absolute equality (45-degree line).
3 With zero representing the global means and one unit increase representing one standard deviation decrease in the perception of corruption.

Descriptive Statistics

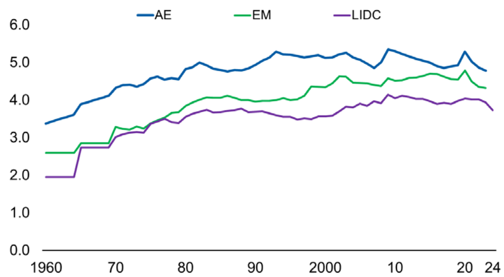
The main advantage of this dataset, particularly for public spending, is that it covers a large set of countries and years by integrating various sources, as described in the previous sections. While the availability and quality of data may differ across indicators and years (see Online Annex Figure.1.1.1), long times series are better suited for empirical analyses that aim to capture the long-term macroeconomic implications of public spending developments.

Online Annex Figure 1.1.1 Spliced Public Spending Variables

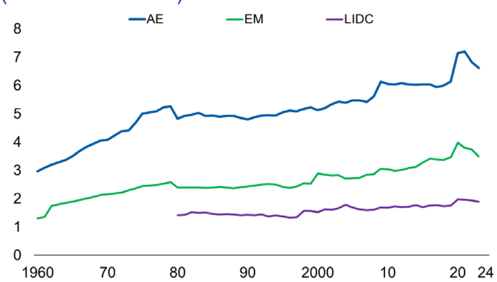
1. Investment
(Percent of GDP)



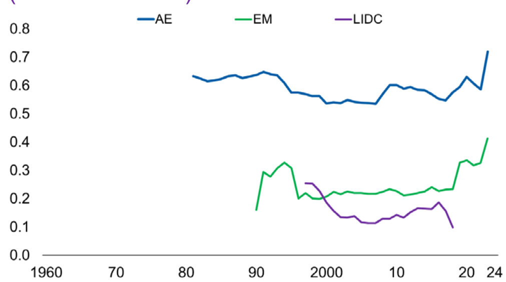
2. Education
(Percent of GDP)



3. Health
(Percent of GDP)



4. R&D
(Percent of GDP)

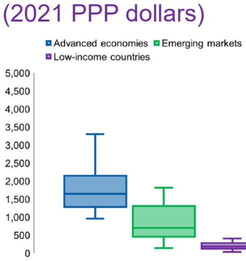


Source: IMF staff calculations.
Note: See Annex Table 1.1 for more information. Using simple unweighted averages.

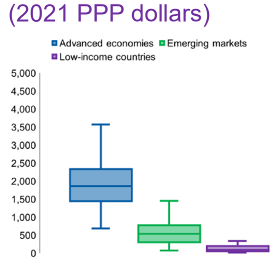
Another advantage is that it covers multiple outputs (or outcome) variables across key areas of public spending which helps capture gaps in efficiencies. Online Annex Figures 1.1.2 and 1.1.3 show that the patterns in efficiency gaps revealed in the chapter are reflected in underlying input and output variables.

Online Annex Figure 1.1.2 Input Variables for Efficiency Estimates

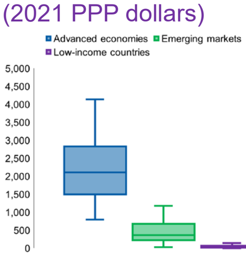
1. Per-capita public investment
(2021 PPP dollars)



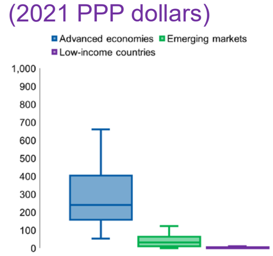
2. Per-capita public education spending
(2021 PPP dollars)



3. Per-capita public health spending
(2021 PPP dollars)



4. Per-capita public R&D spending
(2021 PPP dollars)



Source: IMF staff calculations.
Note: See Annex Table 1.1 for more information. Using country averages over 1980-2024. The horizontal line inside the box shows the median, the box shows the interquartile range, and the whiskers show the smallest and largest data points within 1.5 times the interquartile range from the quartiles (i.e., minimum and maximum values that are not considered outliers).

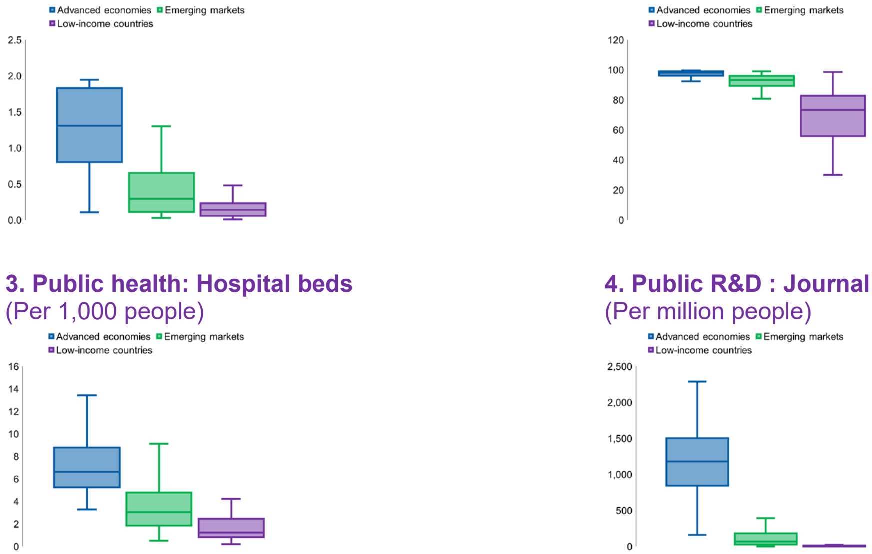
Online Annex Figure 1.1.3 Select Output Variables for Efficiency Estimates

1. Public investment: Road density
(km per square km)

2. Public education: Primary enrollment (net)
(Percent of relevant age groups)

3. Public health: Hospital beds
(Per 1,000 people)

4. Public R&D : Journal articles
(Per million people)



Source: IMF staff calculations.
Note: See Annex Table 1.2 for more information. Using country averages over 1980-2024. The horizontal line inside the box shows the median, the box shows the interquartile range, and the whiskers show the smallest and largest data points within 1.5 times the interquartile range from the quartiles (i.e., minimum and maximum values that are not considered outliers).

The database also captures various drivers of public spending developments. The correlation matrix of selected drivers is reported in Online Annex Table 1.1.3 and suggests that many are correlated with each other. This motivates the robustness checks that use variation between and across countries, control for GDP per capita, and control for other combinations of variables in the Bayesian Model Averaging (Online Annex 1.4).

Online Annex Table 1.1.3 Correlation Matrix for Selected Drivers of Public Spending Development

Variables	(1)	(2)	(3)	(4)	(5)
(1) Expenditure decentralization	1.000				
(2) Control of corruption	0.440***	1.000			
(3) Gross government debt	0.202***	0.036**	1.000		
(4) Effectiveness of public investment management (implementation)	-0.018	0.503***	0.008	1.000	
(5) Rule of law	0.199***	0.777***	0.118***	0.325***	1.000

Source: IMF staff calculations.
Note: See Annex Table 1.3 for more information. *** p<0.01, ** p<0.05, * p<0.1.

Online Annex Figure 1.1.4 shows public investment, education, and compensation spending in selected countries that underwent significant reforms, as discussed in the main text.

Online Annex Figure 1.1.4 Select Countries for Public Spending

1. Public investment
(Share of total expenditure)

2. Public education spending
(Share of total expenditure)

Source: IMF World Economic Outlook Database; IMF Government Finance Statistics; IMF FAD Government Compensation and Employment Dataset; IMF staff calculations.
Note: Periods of reforms are highlighted in gray: 2019–24 for Serbia and 2002–2008 for Brazil.

Online Annex 1.2. Public Spending Efficiency Gap Estimates[1]

This Online Annex details the estimation of public spending efficiency gaps presented in the main text. It first outlines key conceptual considerations, followed by a description of the estimation approach. The Online Annex then briefly illustrates the underlying data and presents selected results on time-varying efficiency gaps across government investment, healthcare, education, and R&D spending for 174 countries over 1980–2023.

Conceptual Considerations

Public spending efficiency, or technical efficiency, is defined as how close governments come to achieving the maximum possible outputs (or outcomes) given a fixed level of inputs (public expenditure). Efficiency estimates measure the gap between observed outcomes and those achievable under best-practice management, technology, and institutional arrangements.

The measurement of efficiency gaps across a large sample of countries poses several significant challenges:

- First, a critical challenge is separating controllable inefficiency from external factors beyond government control. Countries may exhibit certain outcomes due to external shocks or circumstances beyond managerial control.
- Second, the determination of public services is complex due to its inherent multifaceted nature. Infrastructure spending, for instance, simultaneously produces transportation capacity, electricity access, telecommunications connectivity, and water services. These multiple outputs complicate simple aggregation.
- Third, many public services are inherently hard to measure consistently across countries. For instance, the efficient provision of reliable electricity access typically involves establishing a high-quality grid. While this could, in principle, be measured by total installed electricity generation capacity (in gigawatts per capita), such data are not available for a sufficiently large and comparable set of countries.
- Fourth, efficiency estimates for a large sample of countries ought to reflect the fact that countries are markedly different, most notably in terms of their development stage, geographical characteristics, and the extent of private spending. Therefore, applying the same production function could attribute many country differences to inefficiency estimates.

The framework presented in the chapter addresses these challenges as follows:

- To distinguish persistent inefficiency from random shocks or transitory factors beyond the government's control, the analysis relies on stochastic frontier methods. These methods allow for observed deviations from potential performance to be decomposed into inefficiency and noise components. [2] Moreover, uncertainty regarding the specification of the production function is addressed using model-averaging techniques.[3]

[1] Prepared by Krzysztof Bańkowski, Chloe Hyungsun Cho, and Galen Sher.

[2] Afonso et al. (2023) provide a comprehensive overview of frontier methods for assessing efficiency, which are broadly classified into two main groups. Non-parametric approaches, such as Data Envelopment Analysis and Free Disposal Hull, do not impose a specific functional form and instead construct the frontier from observed data using mathematical programming. In contrast, parametric approaches, like Stochastic Frontier Analysis and Corrected Ordinary Least Squares (COLS), presuppose a particular functional form for the frontier and estimate its parameters using econometric techniques.

[3] A primary limitation of the SFA methodology is its requirement to assume a specific functional form for the production frontier, which introduces a risk of model misspecification. The analysis in the October 2025 Fiscal Monitor mitigates this issue by adopting a flexible *translog* distance function and estimating a comprehensive set of specifications corresponding to all non-empty subsets of the considered variables.

- To reflect the multifaceted nature of public service delivery, the analysis uses multi-output distance functions that preserve the diversity of government outputs—such as transportation, electricity, and water—instead of collapsing them into a single composite indicator. This allows for a more accurate representation of complex input–output relationships.
- In the absence of harmonized data for some service dimensions, pragmatic proxies are employed. For example, electricity consumption per capita serves as a powerful substitute for unavailable indicators such as total installed generation capacity, effectively capturing the quality and depth of electricity access.
- Country-specific characteristics, like differences in income per person or private spending, are accounted for by fixed effects (Greene 2005) which accommodate both structural differences and time-invariant inefficiencies, while crucially allowing for time-varying inefficiency estimates. This approach ensures that the production function is effectively tailored to each country while maintaining a unified analytical framework. Global time-varying factors, such as technological progress, are captured through time fixed effects.

The efficiency of government spending has been an important theme in literature (Afonso et al. 2023). Notable studies with global coverage include Herrera et al. (2025) and Apeti et al. (2023) which span multiple sectors, as well as sector-specific analyses such as Kapsoli and Mogues (2023) for infrastructure and Garcia-Escribano et al. (2022) for healthcare. Studies consistently find substantial efficiency gaps, suggesting considerable scope for improving the efficiency of public expenditures in many countries. The chapter contributes to this literature in three important ways. First, it provides time-varying estimates. Second, it adopts multi-output distance functions, which enable us to capture the multi-dimensional nature of public services while avoiding arbitrary aggregation into a single composite measure. Third, the analysis addresses model uncertainty by employing model-averaging techniques.

Econometric Framework

Stochastic Frontier Analysis

The efficiency gap estimation is grounded in the Stochastic Frontier Analysis (SFA) framework originally developed by Aigner, Lovell, and Schmidt (1977), subsequently extended to accommodate multiple outputs through distance function approaches. The fundamental stochastic production relationship is specified as:

$$y_{it} = f(x_{it};\beta) \cdot \exp(v_{it} - u_{it}) \qquad \text{(Eq.2.1)}$$

where $\mathrm{u_{it}} \geq 0$ denotes technical inefficiency, and $\mathrm{v_{it}} \sim \mathrm{N}(0, \sigma_v^2)$ captures random statistical noise for country i at time t.[4]

Estimable equation

Building on the above theoretical foundation, several steps are required to derive the estimable equation used in the regressions.

- First, the distance of observed output from the frontier, $\frac{y_{it}}{f(x_{it};\beta)}$, can be expressed using the distance function $D_0(x_{it}, y_{it}; \beta)$, leading to the following relationship: $1 = D_0(x_{it}, y_{it}; \beta) \cdot \exp\{u_{it} - v_{it}\}$.

[4] The composite error structure $\varepsilon_{it} = \mathrm{v_{it}} - \mathrm{u_{it}}$ allows decomposition of deviations from the frontier into controllable inefficiency and uncontrollable external factors. The split of the error constitutes the major advantage of the SFA method.

- Second, noting that the distance function is homogeneous of degree +1 in outputs allows normalization by the Euclidean norm $|y_{it}|$, yielding: $|y_{it}|^{-1} = D_0\left(x_{it}, \frac{y_{it}}{|y_{it}|}; \beta\right) \cdot \exp\{u_{it} - v_{it}\}$.[5]
- Third, taking logarithms results in the estimable equation used in the regression, where $\ln\big(D_0(\cdot)\big)$ is specified as a *translog* function: $-\ln(|y_{it}|) = \ln\left(D_0\left(x_{it}, \frac{y_{it}}{|y_{it}|}; \beta\right)\right) + u_{it} - v_{it}$.[6]
- Finally, supplementing the equation with country and time-fixed effects (μ_i and ϕ_t respectively) brings the estimable equation below, with predicted u_{it} being the main object of interest. For the estimation, the positive inefficiency term u_{it} is assumed to follow an exponential distribution, which is computationally more stable than alternative choices. Country fixed effects account for unobserved, time-invariant heterogeneity across countries. Their treatment is discussed further below. Time fixed effects, in turn, control for global shocks and common temporal trends that systematically affect all countries.

$$-\ln(|y_{it}|) = \ln D_0\,(x_{it}, y_{it}/|y_{it}|; \beta) + \mu_i + \phi_t + u_{it} - v_{it} \qquad \text{(Eq.2.2)}$$

Model averaging

The inefficiency estimates u_{it} are based on multiple models specified through the flexible formulation of the *translog* distance function for all possible non-empty sets of variables to ensure robustness. The model results are retained only if they satisfy convergence criteria and identify significant inefficiency (likelihood ratio test p-value < 0.05 for $H_0\colon \sigma_u^2 = 0$).[7]

Final efficiency estimates employ a robust averaging procedure to minimize sensitivity to model specification:

- Individual Model Estimation: Each valid model j produces inefficiency estimates $u_{it}^{(j)}$
- First-Differencing: Compute: $\Delta u_{it}^{(j)} = u_{it}^{(j)} - u_{i,t-1}^{(j)}$ to remove level effects
- Cross-Model Median: Aggregate estimate is $\Delta u_{it} = \text{median}_j\left(\Delta u_{it}^{(j)}\right)$
- Level Reconstruction: $u_{it} = u_{i,t-1} + \Delta u_{it}$.[8]

This procedure reduces sensitivity to extreme model specifications while preserving temporal variation. The use of first differences ensures that the model aggregate is robust to compositional changes in the dataset, such as the introduction or discontinuation of an outcome variable.

Model averaging is performed not only across different combinations of outcome variables j but also across two distinct treatments of country-fixed effects (μ_i). These fixed effects may reflect both persistent

---

[5] The Euclidean norm for M outputs is calculated as follows: $|y_{it}| = \left(\sum_{m=1}^{M} y_{mit}^2\right)^{1/2}$.

[6] For a production technology with one input and M outputs, the *translog* functional form is written as: $\ln\big(D_O(x, y)\big) = \alpha_0 + \sum_{i=1}^{M} \alpha_i \ln y_i + \beta_1 \ln x + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_{ij} \ln y_i \ln y_j + \frac{1}{2}\delta_{11}[\ln x]^2 + \sum_{i=1}^{M} \gamma_{i1} \ln y_i \ln x$.

[7] The number of estimated models depends on the number of potential output variables (M) considered for each category, following the formula $2^{\mathrm{M}} - 1$. Given the variable set, the estimates are based on 511 (M=9) models for public investment spending, 31 (M=5) models for healthcare spending, 31 (M=5) models for education spending, and 15 (M=4) models for R&D spending. The overwhelming majority of the models identify statistically significant inefficiency.

[8] To reconstruct the level series, the 2007–2017 period—characterized by the highest data quality and coverage—is used. The median level over this period is assigned as the anchor value for 2012, the midpoint of the interval. The series is then chain-linked backward and forward using the corresponding median year-on-year changes.

inefficiency—such as those stemming from institutional quality—and structural factors unrelated to efficiency, such as geography or private sector spending. To address this conceptual ambiguity, inefficiency is computed under two methods and then subsequently averaged. The first method yields the estimate u_{it}, which effectively treats the country-fixed effect as part of the production function. The second method treats the fixed effect as a component of inefficiency, yielding the following estimate:

$$w_{it}^{(j)} = \mu_i^{(j)*} + u_{it}^{(j)}, \tag{Eq.2.3}$$

In this equation, $\mu_i^{(j)*} = \mu_i^{(j)} - \min_i \mu_i^{(j)}$ represents the non-negative inefficiency component derived from the country-fixed effect. These model-specific estimates, $w_{it}^{(j)}$, are then aggregated across models to produce the final inefficiency series, w_{it} , using the same procedure of first-differencing, cross-model median derivation, and level reconstruction as was used for u_{it}.

Efficiency gap definition

The inefficiency terms are used to calculate the efficiency gap, which measures the proportional shortfall from full efficiency. Two sets of estimates are obtained based on the alternative treatments of country-fixed effects:

$$\text{Efficiency Gap}_{it}^{u} = 1 - \exp(-u_{it}) \tag{Eq.2.4}$$

$$\text{Efficiency Gap}_{it}^{w} = 1 - \exp(-w_{it}) \tag{Eq.2.5}$$

The final estimate then aggregates these two by taking their simple arithmetic average:

$$E\text{fficiency Gap}_{it} = (\text{Efficiency Gap}_{it}^{u} + \text{Efficiency Gap}_{it}^{w})/2 \tag{Eq.2.6}$$

This metric ranges from 0 (fully efficient) to 1 (completely inefficient) and represents the percentage by which outputs could be increased with current inputs under best-practice technology, while explicitly accounting for random shocks and measurement errors.

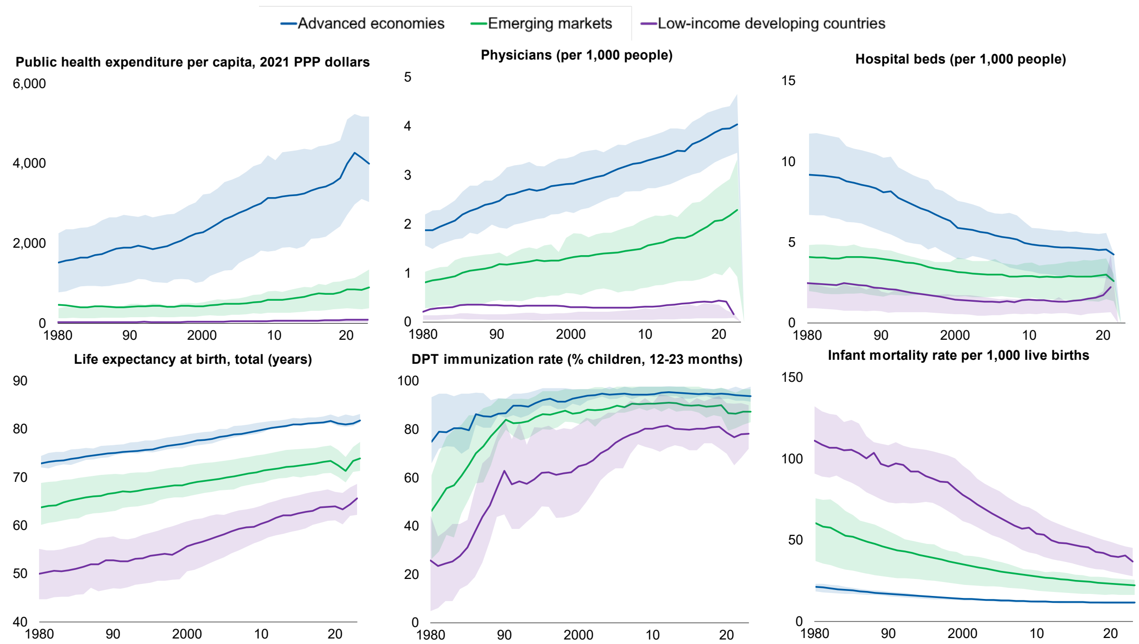
The treatment of fixed effects has a substantial impact on gap estimates. On average, the efficiency gaps are approximately 51 percentage points higher when country-fixed effects are included in the inefficiency term compared to when they are excluded. Consequently, alternative weighting schemes could generate larger (or smaller) efficiency gaps by placing more (or less) weight on the models that include the country-fixed effect in the inefficiency term. In the extreme case, with 100% weight on the latter model, the inefficiency gaps would be 26 percentage points larger than reported in this chapter.

Underlying Data

The input variables used to measure efficiency gaps are functional public expenditures across four key sectors: public investment, health, education and R&D spending. The data are drawn from the database constructed for October 2025 Fiscal Monitor (see Online Annex 1.1 for details). To prepare the variables for regression analysis, some transformations were applied. First, to account for the persistent relationship between spending and outcomes, all public spending series are smoothed using a five-year moving average.[9]

[9] For public investment, related studies—such as Kapsoli and Mogues (2023) and Baum et al. (2020)—typically use the capital stock as the input measure rather than public investment expenditure. In practice, the two measures are closely related, as the capital stock is derived from accumulated spending adjusted for depreciation. Unsurprisingly, the efficiency estimates do not change significantly when capital stock is used instead.

Online Annex Figure 1.2.1 Selected Input and Output Variables for Health Spending Efficiency

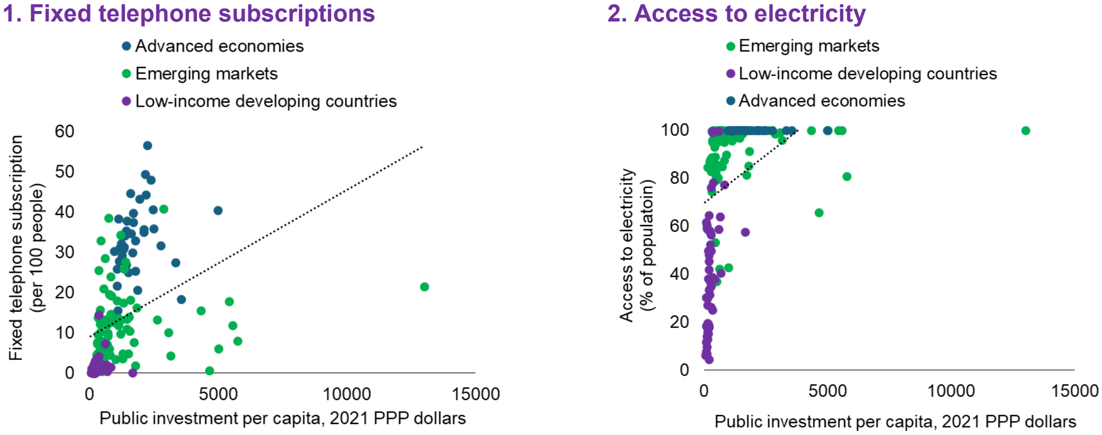


Source: IMF staff calculations.
Note: See Online Annex Tables 1.1.1 and 1.1.2 for details on the variables shown, including data sources. Lines indicate income group means. Dark-shaded areas show interquartile ranges (25th–75th percentiles).

Second, to ensure cross-country comparability, the variables are expressed on a per capita basis and converted to constant 2021 U.S. dollars using purchasing power parity (PPP) adjustors.

Online Annex Figure 1.2.2 Variation of Selected Variables Within Country Groups

1. Fixed telephone subscriptions **2. Access to electricity**



Source: IMF staff calculations.
Note: See Online Annex Tables 1.1.1 and 1.1.2 for details on the variables shown, including data sources. The scatter plot values represent country-level averages over time.

The output variables aim to measure the quantity and quality of public services. These indicators are also drawn from the same database. The selection of output variables was guided by previous studies and constrained by data availability and quality. Some indicators, such as electricity consumption per capita, are included as proxies for government performance, even though they may not directly measure public services and partly reflect private sector contributions. The dataset includes both volume indicators (e.g., consumption levels) and quality measures (e.g., survey data on infrastructure, healthcare and education quality).

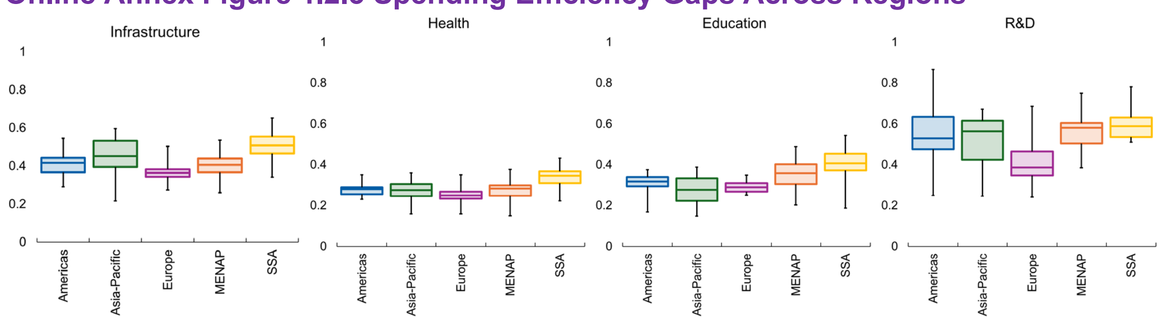
This variable selection strategy aims to maximize cross-country variation while acknowledging the limitations of specific indicators. For instance, while fixed line telephone access varies little among low-income countries, it is a key differentiator for advanced economies (see Online Annex Figure 1.2.1, panel 1). Conversely, access to electricity shows significant variation in low-income countries but not in advanced economies (see Online Annex Figure 1.2.1, panel 2). Including both variables ensures sufficient variation within different country groups. A comprehensive list of outcome variables, with definitions and sources, is provided in Online Annex Table 1.2.1 Using health spending as an example, Online Annex Figure 1.2.2 highlights the temporal evolution of key variables and reveals substantial disparities among different income groups.

Spending Efficiency Gap Estimates

Regional estimates

Beyond presenting efficiency gaps by country groups in the main text, geographic aggregation reveals distinct regional patterns, too (see Online Annex Figure 1.2.3). *Europe* and *the Americas* demonstrate the lowest efficiency gaps on average, with consistently better performance compared to other regions. *African countries* have the largest inefficiencies, consistent with income-based findings. The data also reveal significant differences within, especially for *the Asia-Pacific* and *the Middle East, North Africa, Afghanistan, and Pakistan (MENAP)* regions.[10]

Online Annex Figure 1.2.3 Spending Efficiency Gaps Across Regions



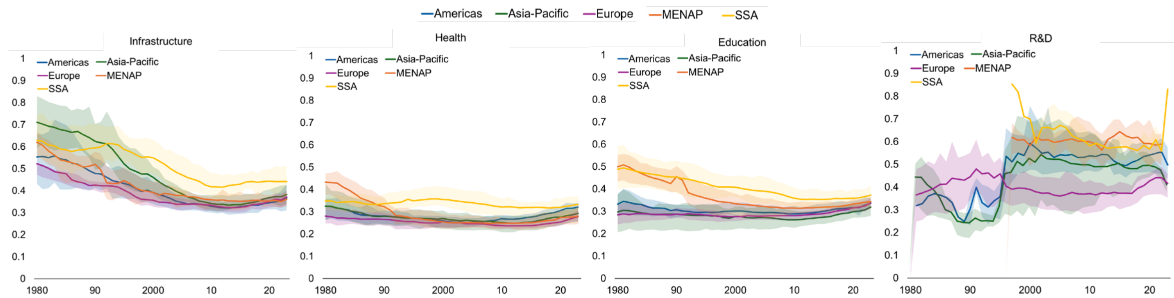
Source: IMF staff estimates.
Note: Efficiency gaps range from 0 (fully efficient) to 1 (fully inefficient). Box plots show regional medians and interquartile ranges (25th–75th percentiles), averaged over time. Whiskers represent the minimum and maximum values.

The estimates also suggest significant regional differences in the evolution of efficiency (see Online Annex Figure 1.2.4). Inefficiencies in *Asia and Pacific and MENAP countries* significantly reduced since the 1980s, especially in the case of physical investment and health spending, reflecting rapid institutional development and capacity building in these regions. Also, *the MENAP region* recorded a notable improvement in the education sector. On the other hand, improvements in *African countries* have not been significant, pointing to persistent structural challenges *Europe* and *the Americas* show relatively stable efficiency levels over time, with usually better performance compared to other regions. Online Annex Figure 1.2.5 shows public

[10] Favorable government spending efficiency gaps within the MENAP region are largely attributable to countries from the Caucasus and Central Asia. This efficiency stems from their ability to deliver relatively positive outcomes despite comparatively low levels of public spending.

investment and education spending efficiency gaps for select countries, as described in the case studies section of the main text.

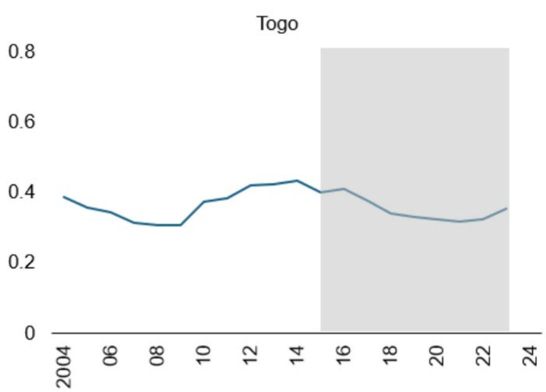
Online Annex Figure 1.2.4 Evolution of Efficiency Gaps Across Regions



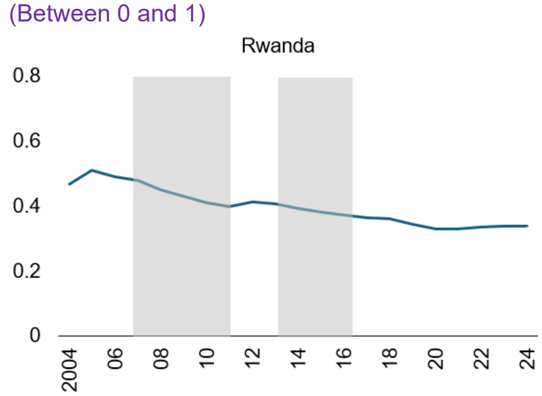
Source: IMF staff estimates.
Note: Efficiency gaps range from 0 (fully efficient) to 1 (fully inefficient). Lines indicate regional group means. Shaded bands show interquartile ranges (25th–75th percentiles).

Online Annex Figure 1.2.5 Select Countries for Efficiency Gaps

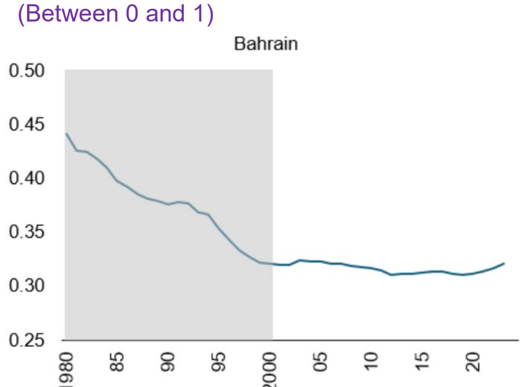
1. Public investment efficiency gap of Togo
(Between 0 and 1)



2. Public education spending efficiency gap of Rwanda
(Between 0 and 1)



3. Public education spending efficiency gap of Bahrain
(Between 0 and 1)



Source: IMF staff estimates.
Note: The charts show select countries for spending efficiency gaps, with periods of notable decreases after reforms highlighted in gray: 2015–2023 for Togo, 2007–2011 and 2013–2016 for Rwanda, and 1980-2000 for Bahrain.

Benchmarking

Online Annex Figure 1.2.5 compares the efficiency gap estimates in this chapter with those reported by Herrera et al. (2025), Kapsoli and Mogues (2023), and Garcia-Escribano et al. (2022). For public investment, the estimates in this chapter are similar to those in the literature. Using a similar country sample and time period as in Herrera et al. (2025), the global average public investment efficiency gap is 0.36 in this study and 0.45 in Herrera et al. (2025). Using a similar country/time sample to Kapsoli and Mogues (2023), the gap is 0.41 in this study and 0.38 in Kapsoli and Mogues (2023). By contrast, the efficiency gaps estimated in this chapter are larger for health and education spending. The comparable health spending gap is 0.27 in this study, compared with 0.13 in Herrera et al. (2025) and 0.07 in Garcia-Escribano et al. (2022). For education spending, the gap is 0.31, compared with 0.22 in Herrera et al. (2025).

The different magnitudes of the efficiency gaps found in this chapter are likely attributable to its methodological innovations. The analytical approach incorporates time-varying estimates, uses fixed effects to control for structural country differences and global technological progress, systematically addresses model uncertainty, and assesses outcomes using a multi-dimensional framework rather than a single arbitrary index. Given these advances, a close alignment with the findings from existing global studies is neither expected nor warranted. Despite differences in magnitude, the results display a strong positive correlation with previous findings, as shown in Online Annex Figure 1.2.6.

Online Annex Figure 1.2.6 Comparison of Efficiency Gap Estimates Against Selected Studies

Source: Garcia-Escribano et al. (2022); Herrera et al. (2025); Kapsoli and Mogues (2023), and IMF staff calculations.
Note: The efficiency gap data from Herrera et al. (2025) were derived by averaging all available output-oriented efficiency scores for the respective sectors. The vertical axes, and "IMF average," show estimates from this chapter of the Fiscal Monitor. The sample of countries and years of each panel matches those used in the respective paper from the literature specified on the horizontal axis.

Online Annex 1.3. Public Spending Rigidity Estimates[1]

This Online Annex describes the methodology for estimating public spending rigidity—the inflexibility of changing spending over time. The rigidity estimates form a novel database covering 151 countries from 2000.

Input Data

The analysis estimates rigidity for various public spending categories: total public expenditure, public investment, and the functional public spending categories ("COFOG"). Total public expenditure is from the IMF's Public Finances in Modern History (PFMH) database which includes debt service costs, while public investment is sourced from the IMF's World Economic Outlook database (public gross fixed capital formation, code NFIG). Additionally, functional classifications from IMF's Government Finance Statistics (GFS) COFOG categorize expenses related to defense, education, health, social protection, and R&D. In this dataset, defense and public order spending is the sum of Expenditure on Defense (GF02) and Expenditure on Public Order & Safety (GF03), with pension spending accounted for through Expenditures on Old Age and Survivors (GF1002, GF1003). R&D components are extracted from Expenditures on Education(GF09), Health(GF07), and Social Protection(GF1009), hence, it does not include R&D spending of defense and public order (which is included in defense and public order). The dataset also includes public sector wage bills from the IMF's Government Compensation and Employment Dataset (IMF 2016). The public wage bill is defined as general government compensation of employees. If general government data are not available, central government compensation data are used. Country coverage for each spending category is outlined in Online Annex Table 1.3.1. The analysis provides public spending rigidity estimates for 151 countries for total public expenditure, 155 countries for public investment spending, 188 countries for public compensation spending, and a lower number of countries—ranging from 50 to 62 for functional categories due to data limitations.

Online Annex Table 1.3.1 Country Coverage

	Advanced Economies	Emerging Markets	Low-Income Developing Countries
Total Public Expenditure	33	77	41
Public Investment	33	79	44
Defense and Public Order	35	21	5
Public Education Spending	35	21	6
Public Health Spending	35	21	6
Social Protection	35	21	2
Public R&D Spending	31	16	6
Pension Spending	31	13	6
Compensation Spending	38	92	58

Source: IMF staff compilation.

Computational Methodology

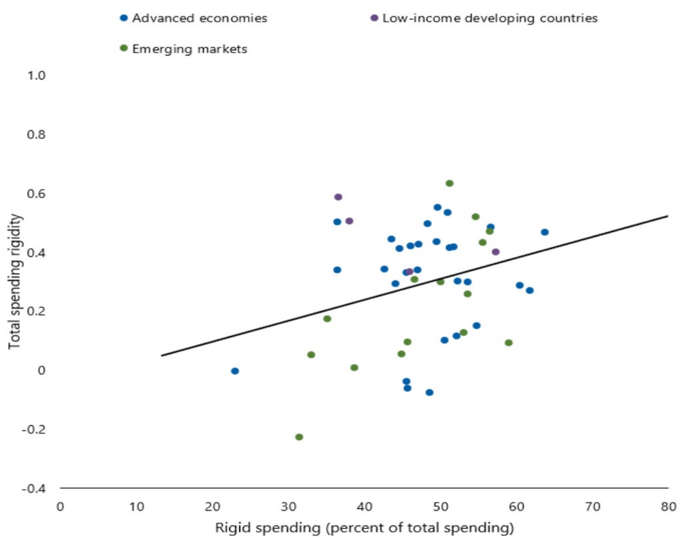
Public spending rigidity is computed as one-year lag autocorrelation, as presented in Piguillem and Riboni (2024). A value of 1 means the highest level of rigidity. The original methodology is extended by employing rolling-windows to create a panel dataset of country-level rigidity estimates which vary over time. A longer rolling window can help mitigate the effects of local shocks but requires higher data quality in specific spending areas. Given the relatively short time series available for emerging market and developing

[1] Prepared by Hongchi Li and Zsuzsa Munkacsi.

economies, the analysis employs five-year rolling windows. For instance, the rigidity estimate for 2020 is derived from the correlation between spending series between 2015-2019 and 2016-2020.

The estimates are positively correlated with measures obtained in the literature through other methods—such as the sum of public compensation, pensions and interest payments (Herrera and Olaberria 2020). As shown in Online Annex Figure 1.3.1, advanced economies exhibit higher levels of rigid spending and rigidity of total public spending compared to emerging markets and low-income developing countries, suggesting that larger governments with nondiscretionary spending encounter more rigid budget constraints (Mattina and Gunnarsson 2007).

Online Annex Figure 1.3.1 Rigidity Estimates of Public Spending



Source: PFMH; GFS; WEO; and IMF staff calculations.
Note: For each country, rigidity is calculated as the one-year autocorrelation of each spending category as a share of total expenditure, in rolling windows of five years. Rigid spending is the sum of public compensation, pensions, and interest payments following Herrera and Olaberria (2020).

Descriptive Statistics

Online Annex Tables 1.3.2 and 1.3.3 present key statistical measures of the dataset categorized by country groups. For each spending category, rigidity estimates are expressed as share of total spending and as share of GDP. It is shown that, on average, advanced economies exhibit greater rigidity in spending across most categories, particularly in education, health and social protection. Conversely, low-income developing countries show comparatively lower rigidities in most spending categories, although the differences are less sizeable in the cases of public investment and compensation. Furthermore, the time-varying estimates derived from a rolling window approach enable an analysis of how rigidity changes over time. As shown in Online Annex Figure 1.3.2, advanced economies are experiencing increasingly rigid spending across all spending categories. At the same time, in emerging markets spending rigidity has declined in many areas such as defense and public order, education, and health.

Online Annex Table 1.3.2 Summary Statistics by Country Groups
(Percent of expenditure)

		Advanced Economies	Emerging Markets (excl. China)	Low-Income Developing Countries
Public Investment	Mean	0.26	0.29	0.24
	Std. Dev	0.49	0.49	0.48
	25th pct	-0.09	-0.04	-0.10
	Median	0.31	0.35	0.26
	75th	0.67	0.72	0.64
Defense and Public Order	Mean	0.24	0.31	0.12
	Std. Dev	0.48	0.43	0.47
	25th pct	-0.12	-0.02	-0.18
	Median	0.26	0.40	0.10
	75th	0.67	0.68	0.43
Public Education Spending	Mean	0.28	0.13	0.16
	Std. Dev	0.48	0.45	0.53
	25th pct	-0.07	-0.22	-0.34
	Median	0.35	0.10	0.23
	75th	0.69	0.52	0.66
Public Health Spending	Mean	0.37	0.23	-0.16
	Std. Dev	0.50	0.49	0.46
	25th pct	-0.02	-0.11	-0.62
	Median	0.49	0.28	-0.11
	75th	0.81	0.63	0.21
Social Protection	Mean	0.32	0.29	-0.23
	Std. Dev	0.49	0.48	0.35
	25th pct	-0.06	-0.05	-0.56
	Median	0.43	0.33	-0.29
	75th	0.72	0.72	0.16
Public R and D Spending	Mean	0.29	0.22	0.13
	Std. Dev	0.48	0.47	0.51
	25th pct	-0.09	-0.20	-0.26
	Median	0.34	0.29	0.17
	75th	0.70	0.62	0.58
Pension Spending	Mean	0.30	0.28	0.28
	Std. Dev	0.49	0.48	0.47
	25th pct	-0.11	-0.04	-0.04
	Median	0.38	0.34	0.33
	75th	0.74	0.68	0.67
Compensation Spending	Mean	0.32	0.33	0.35
	Std. Dev	0.48	0.44	0.34
	25th pct	-0.06	0.04	0.16
	Median	0.33	0.44	0.30
	75th	0.77	0.67	0.56

Source: IMF staff calculations.

Online Annex Table 1.3.3 Summary Statistics by Country Groups

(Percent of GDP)

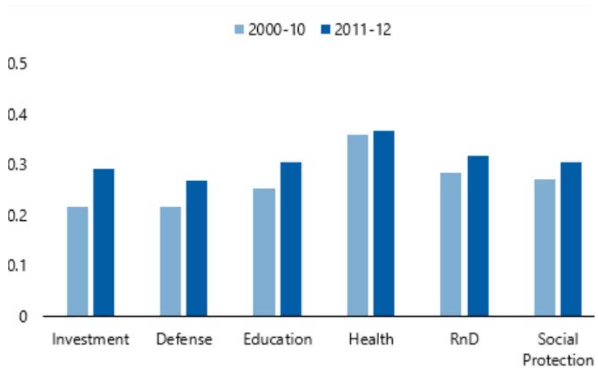
		Advanced Economies	Emerging Markets (excl. China)	Low-Income Developing Countries
Total Public Expenditure	Mean	0.36	0.25	0.29
	Std. Dev	0.45	0.46	0.46
	25th pct	0.05	-0.08	-0.05
	Median	0.41	0.28	0.32
	75th	0.75	0.64	0.68
Public Investment	Mean	0.29	0.30	0.26
	Std. Dev	0.46	0.47	0.47
	25th pct	-0.06	-0.04	-0.10
	Median	0.33	0.33	0.30
	75th	0.69	0.71	0.66
Defense and Public Order	Mean	0.33	0.31	0.14
	Std. Dev	0.47	0.42	0.46
	25th pct	0.01	0.02	-0.10
	Median	0.40	0.32	0.20
	75th	0.75	0.65	0.42
Public Education Spending	Mean	0.37	0.13	0.35
	Std. Dev	0.44	0.45	0.38
	25th pct	0.07	-0.26	0.07
	Median	0.44	0.13	0.41
	75th	0.73	0.50	0.65
Public Health Spending	Mean	0.43	0.21	0.08
	Std. Dev	0.43	0.49	0.50
	25th pct	0.14	-0.14	-0.20
	Median	0.53	0.27	0.03
	75th	0.78	0.56	0.45
Social Protection	Mean	0.41	0.24	0.23
	Std. Dev	0.44	0.45	0.52
	25th pct	0.09	-0.12	-0.26
	Median	0.50	0.27	0.40
	75th	0.79	0.60	0.63
Public R and D Spending	Mean	0.32	0.17	0.37
	Std. Dev	0.47	0.52	0.50
	25th pct	-0.04	-0.23	0.43
	Median	0.41	0.22	0.53
	75th	0.72	0.67	0.67
Pension Spending	Mean	0.45	0.40	0.21
	Std. Dev	0.44	0.46	0.64
	25th pct	0.11	0.13	-0.30
	Median	0.57	0.50	0.40
	75th	0.84	0.80	0.63
Compensation spending	Mean	0.40	0.33	0.38
	Std. Dev	0.45	0.45	0.49
	25th pct	0.08	0.00	0.05
	Median	0.48	0.39	0.49
	75th	0.81	0.72	0.79

Source: IMF staff calculations.

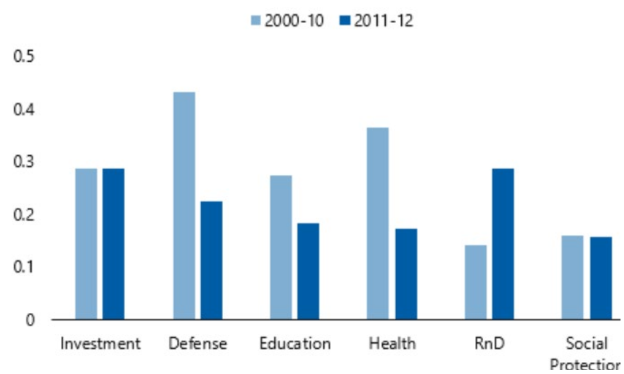
Online Annex Figure 1.3.2 Rigidity Estimates Over Time

(Between 0 and 1)

1. Advanced economies



2. Emerging market economies



Source: GFS; WEO; and IMF staff calculations.
Note: The figures show the average of country-level rigidity of each spending category during 2000-10 and 2011-22. Rigidity is calculated as the one-year autocorrelation of each spending category as a share of total expenditure, in rolling windows of five years.1 (0) denotes maximum (minimum) rigidity.

Online Annex 1.4. Methodology for Identifying Determinants of Public Spending Developments[1]

This Online Annex details the methodology used to analyze the drivers of public spending efficiency, composition and rigidity. It presents selected results that are not shown in the main text and tests the results for robustness.

Data and Variable Construction

The analysis focuses on four key pro-growth public expenditure categories—investment, education, health, and R&D—examining both spending efficiency gaps and their shares as a percent of total public spending. Explanatory variables are grouped into three broad categories: cyclical factors (e.g., crises, elections), structural characteristics (e.g., demographics, political ideology), and institutional-policy drivers (e.g., decentralization, government quality, fiscal rules). Except for categorical or dummy variables, these variables are standardized—the sample-wide mean is subtracted, and the difference is divided by the standard deviation— to ensure interpretability of results. Regression coefficients thus measure the association of a one standard deviation increase in the driver with the change in the outcome variable.

The analysis relies on a harmonized panel dataset constructed by merging a wide range of cross-country sources (see Online Annex 1.1). The merged dataset spans over 194 countries and covers the period 1960-2024. Some variables are trimmed by removing outliers, and missing data are addressed by interpolating values for structural variables that are not updated annually, where appropriate. The final panel is unbalanced, reflecting the varying availability of data across countries and years, but retains a broad and representative sample.

Empirical Methodology

Statistically significant determinants are identified by applying four methodologies. First, univariate cross-sectional relationships are explored using data average across time. Second, univariate country-year panel relationships are explored with and without fixed effects. Third, robustness is investigated by controlling GDP per person, and fourth, by using a full Bayesian Model Averaging (BMA) procedure.

Cross-sectional relationships

The baseline approach uses ordinary least squares (OLS) regressions of average efficiency gaps scores or spending shares on individual drivers. Country-level data are averaged over 2000–24. Model equation is

$$\overline{y}_i = \alpha + \beta \overline{x}_i + \varepsilon_i \qquad \text{(A4.1)}$$

and standard errors are computed using the heteroskedasticity-consistent method proposed by MacKinnon and White 1985, which adjusts for non-constant variance in the error terms, improving reliability of hypothesis testing across all regressions. Thus, the coefficient β measures the cross-sectional effect of each driver over the past 25 years.

Panel relationships

To account for unobserved heterogeneity, the same relationships are re-estimated using annual panel data with and without country and year fixed effects. For the no-FE model the functional form is

[1] Prepared by Danila Smirnov.

$$y_{it} = \alpha + \beta x_{it} + u_{it} \tag{A4.2}$$

while the two-way FE model augments the right-hand side with C-1 country and T-1 year dummies. Inference uses Driscoll–Kraay standard errors that are robust to heteroskedasticity, serial correlation, and spatial dependence. For variables that are strictly country invariant (e.g., fixed rules) perfect collinearity with dummies is removed via QR decomposition so that all reported slopes come from a full rank design matrix.

Controlling for GDP per person

Additional robustness check isolates relationships between outcome and policy variables from income effects. This is done by residualizing both dependent and explanatory variables on log GDP per person, using the Frisch-Waugh-Lovell decomposition: first, regressing the outcome y_{it} on a constant and log(GDP per person) and save the residuals $\hat{\varepsilon}_{it}^{y}$; second, regressing the policy variable x_{it} on the same controls and save the residuals $\hat{\varepsilon}_{it}^{x}$, and, third, estimating the bivariate model of

$$\hat{\varepsilon}_{it}^{y} = \alpha + \beta \hat{\varepsilon}_{it}^{x} + u_{it} \tag{A4.3}$$

A similar three-step procedure is applied to the cross-section, the plain panel, and the two-way fixed effects panel. Results therefore report slope coefficients β that measure the association between the driver and the efficiency (or composition) outcome conditional on a constant level of GDP per person.

Bayesian Model Averaging

To test robustness to multiple controls, Bayesian Model Averaging (BMA) is used.[2] Three BMA exercises are performed for every dependent variable: plain panel without FEs; panel with country- and year-fixed effects forced into every model; and 2000-24 cross-section. Each chain draws 100,000 models after a burn-in of 10 sweeps with a uniform model prior and a dilution prior that sets the expected model size to one variable. Fixed effects are handled so that posterior probabilities apply only to substantive regressors. For every regressor posterior mean, posterior standard deviation, and posterior inclusion probability (PIP)—which equals the share of visited models in which the variable enters—are analyzed. Variables with posterior t-statistics above 1 in absolute value are interpreted as robust drivers (Steel 2020).

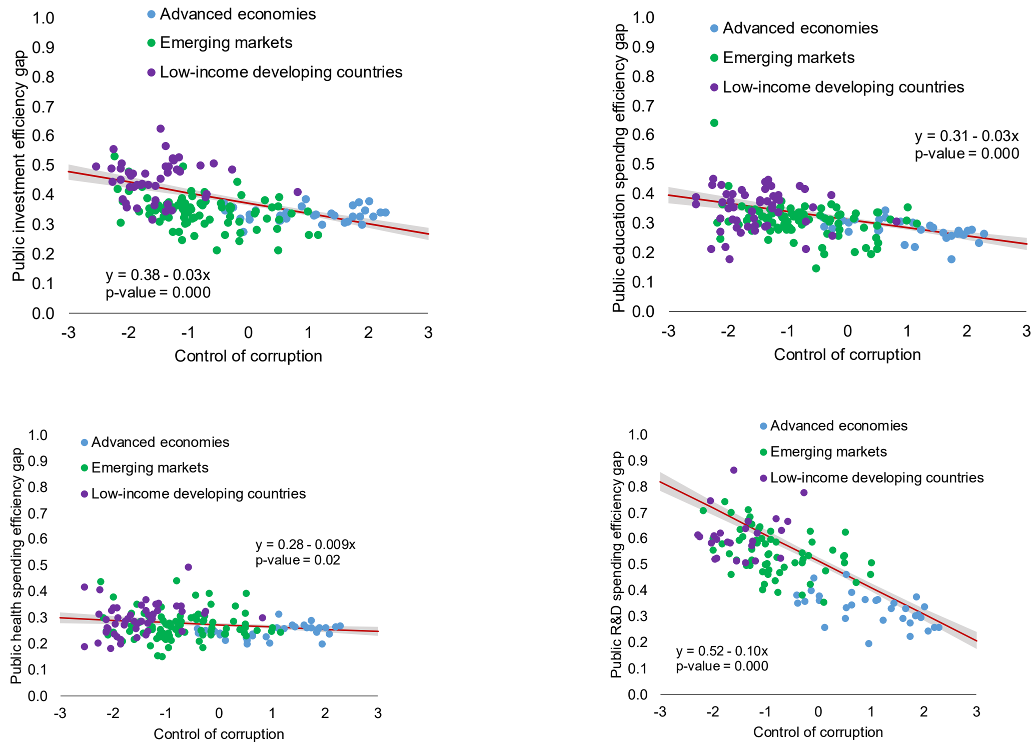
Results

Institutional Quality

Figure 1.4 Panel 1 in the main text shows the negative association between control of corruption and efficiency gap of public investment. Online Annex Figure 1.4.1 shows that these relationships hold also for other categories of pro-growth spending—health, education, and R&D. Online Annex Table 1.4.1 shows that these relationships are robust to controlling for GDP per person for education and R&D and robust to controlling for country and year fixed effects for investment, education, and R&D.

[2] The BMA searches thousands of model combinations and calculates posterior inclusion probabilities and weighted-average coefficients and standard errors for each driver.

Online Annex Figure 1.4.1 Control of Corruption and Efficiency



Source: Worldwide Governance Indicators database; and IMF staff calculations.
Note: Country averages are used for plotting; regression lines reflect cross-section OLS estimates.

Online Annex Table 1.4.1 Public-Spending Efficiency—Significance of Key Institutional Drivers

Institutional Variable	Sector	XS	FE	GDP-Ctrl
Control of corruption	Investment	Y	Y	N
	Education	Y	Y	Y
	Health	Y	N	N
	R&D	Y	Y	Y
Rule of law	Investment	Y	Y	Y
	Education	Y	Y	Y
	Health	Y	N	N
	R&D	Y	Y	Y
State fragility	Investment	Y	Y	Y
	Education	Y	Y	Y
	Health	Y	Y	Y
	R&D	Y	Y	Y
Decentralisation	Investment	Y	Y	N
	Education	Y	Y	N
	Health	N	N	N
	R&D	Y	Y	N

Source: IMF staff estimates.
Note: Y = positive and significant at 10 percent or better; N = not significant. XS = simple cross-section; FE = country + year fixed effects; GDP-Ctrl = controlling for GDP per person.

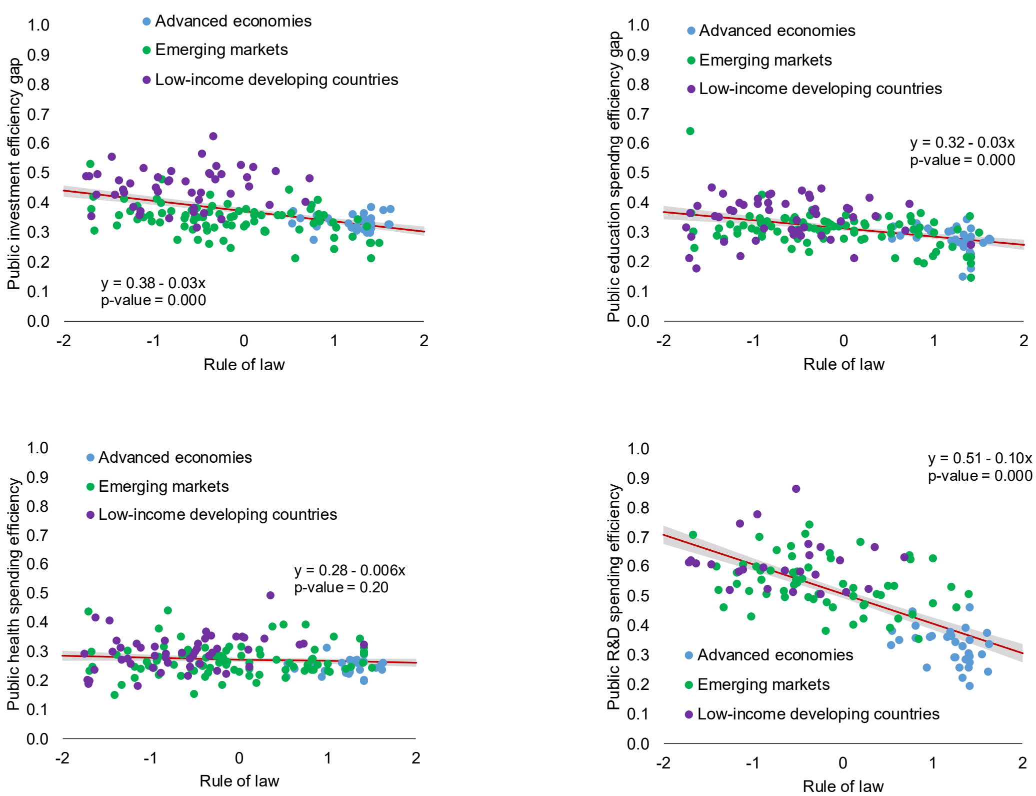
Similarly, the main text notes that rule of law is negatively associated with spending efficiency gaps. Online Annex Figure 1.4.2 demonstrates that this association holds for three categories of pro-growth spending—

investment, education, and R&D. Online Annex Table 1.4.1 shows that this relationship is robust to controlling for GDP per person and country and year fixed effects in all three spending areas.

As discussed in the main text, better public investment management practices, especially the effectiveness of resource allocation, are also related to more efficient spending. Online Annex Figure 1.4.3 illustrates that this relationship holds for public investment, education, and R&D spending efficiency.

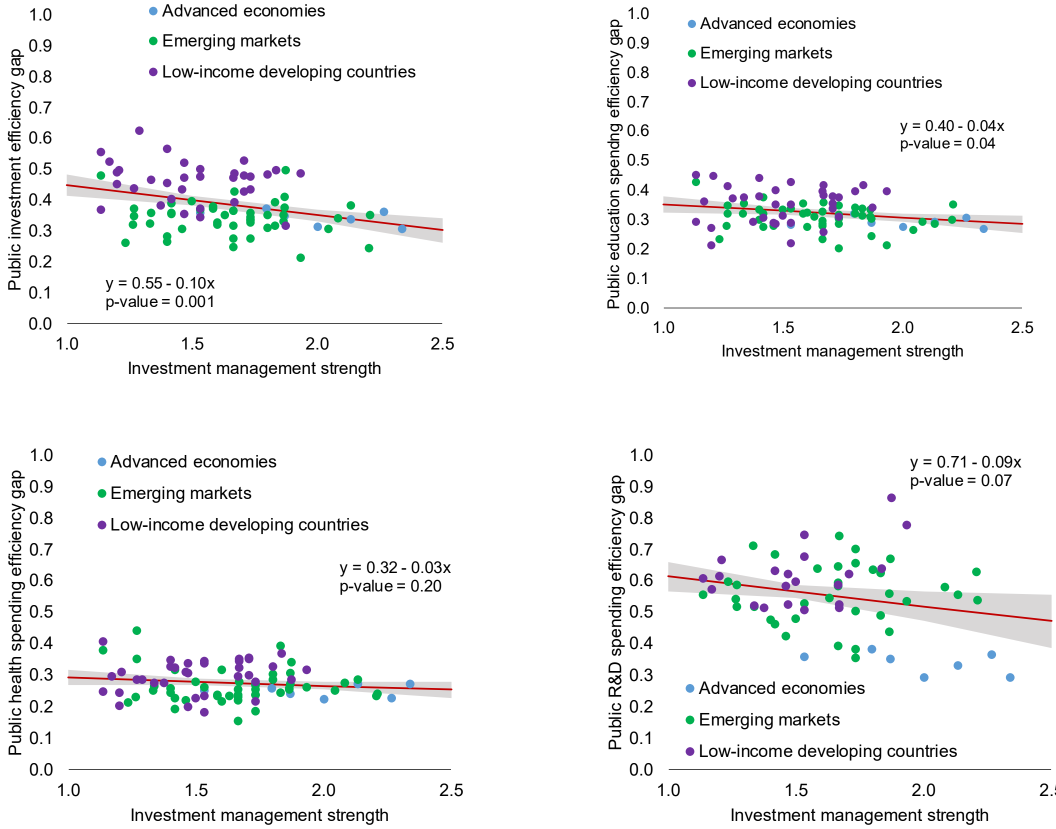
Fragile and conflict-affected states tend to exhibit significantly larger efficiency gaps in pro-growth spending. This is likely due to weaker institutional capacity, governance challenges, and limited absorptive capacity. Online Annex Figure 1.4.4 shows that fragile countries have lower efficiency in public investment, education, health, and R&D spending compared to non-fragile countries. Online Annex Table 1.4.1 indicates that these results are robust to controlling for country and year fixed effects and GDP per person. These findings are also consistent with broader evidence that fragility amplifies the negative impact of external shocks and undermines development outcomes (Jaramillo et al. 2023).

Online Annex Figure 1.4.2 Rule of Law and Efficiency



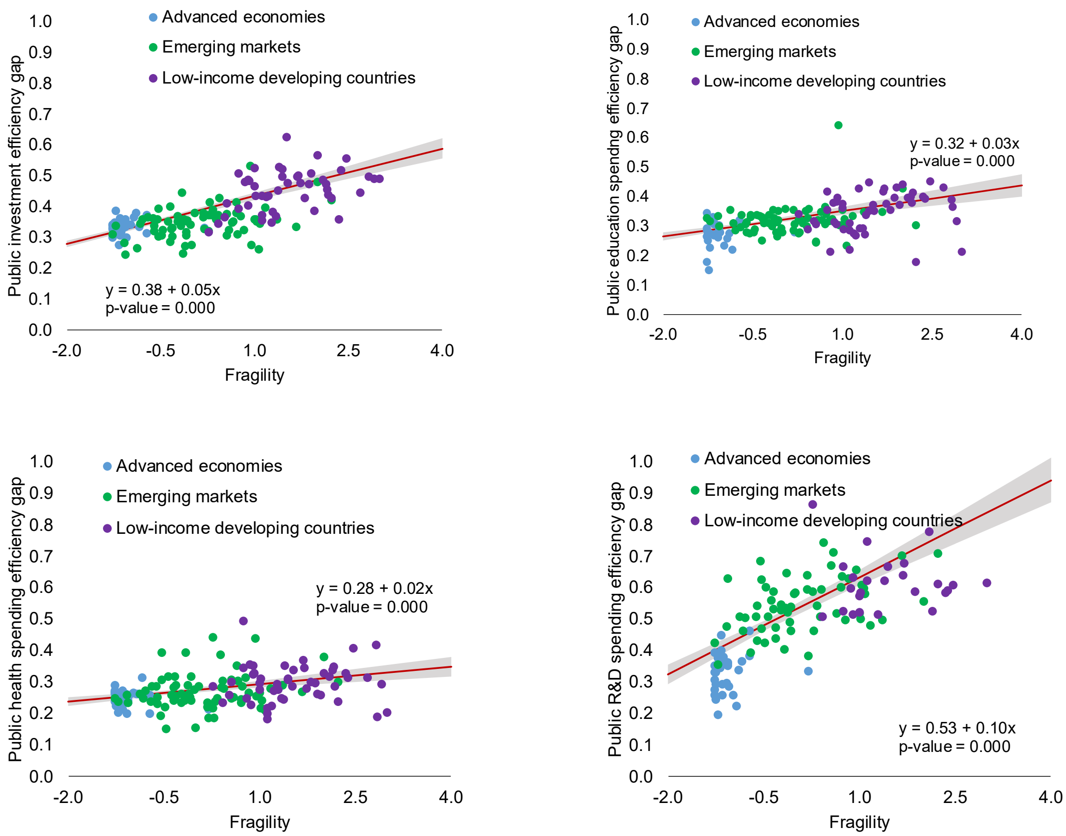
Source: Quality of Government database; and IMF staff estimates.
Note: Country averages are used for plotting; regression lines reflect cross-section OLS estimates.

Online Annex Figure 1.4.3 Public Investment Management and Efficiency



Source: Public Investment Management Assessment scores; and IMF staff estimates.
Note: Higher scores indicate stronger practices. Scores are normalized. Country averages are used for plotting; regression lines reflect cross-section OLS estimates.

Online Annex Figure 1.4.4 State Fragility and Efficiency

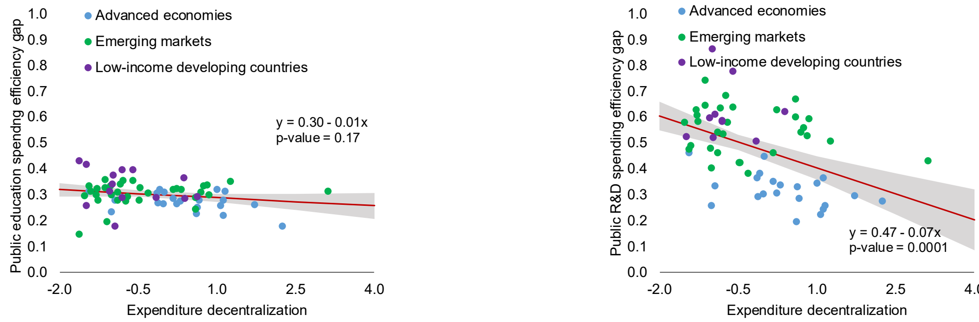


Source: Worldwide Governance Indicators database; and IMF staff estimates.
Note: Higher scores indicate higher fragility. Scores are normalized. Country averages are used for plotting; regression lines reflect cross-section OLS estimates.

Fiscal Architecture

The main text notes that decentralization of spending is associated with narrower efficiency gaps, especially for public education and R&D spending (see Online Annex Figure 1.4.5). Online Annex Table 1.4.1 shows that this relationship holds also for investment spending and is robust to controlling for country fixed effects.

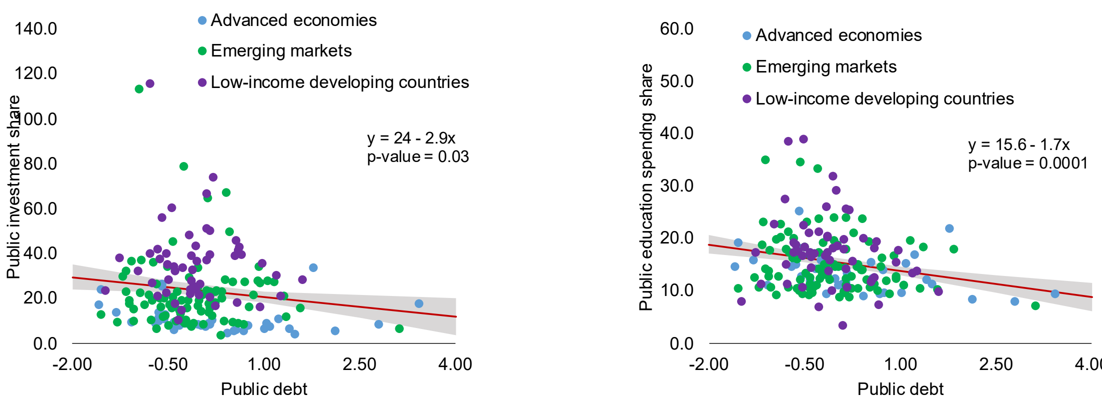
Online Annex Figure 1.4.5 Decentralization and Efficiency



Source: IMF Fiscal Decentralization database; and IMF staff estimates.
Note: Country averages are used for plotting; regression lines reflect cross-section OLS estimates.

The main text also notes that lower public debt is associated with higher allocations towards pro-growth spending. Online Annex Figure 1.4.6 shows that these relationships hold for spending on investment and education. Online Annex Table 1.4.2 indicates that these results are robust to controlling for country and year fixed effects and GDP per person.

Online Annex Figure 1.4.6 Public Debt and Composition



Source: WEO database and IMF staff estimates.
Note: General government gross debt is expressed as a percent of GDP and normalized. Composition refers to the share of spending allocated to pro-growth sectors. Country averages are used for plotting; regression lines reflect cross-section OLS estimates.

Online Annex Table 1.4.2 Spending Composition—Effect of Lower Public Debt

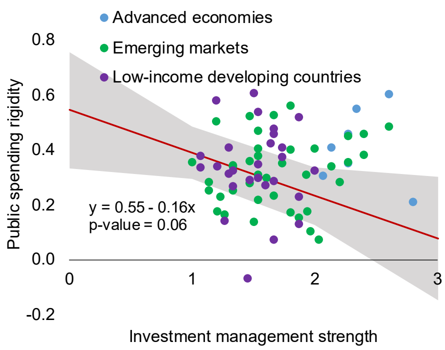
Outcome (Spending Share)	XS	FE	GDP-Ctrl
Investment	Y	Y	Y
Education	Y	Y	Y
Health	N	Y	Y
R&D	N	N	N

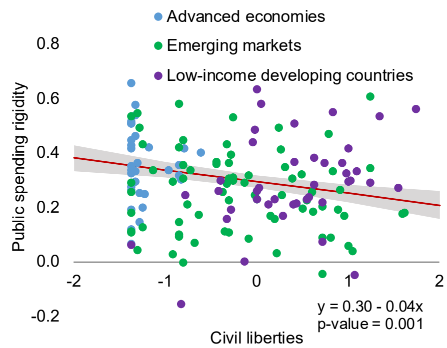
Note: Y = positive and significant at 10 percent or better; N = not significant. XS = simple cross-section; FE = country + year fixed effects; GDP-Ctrl = two-stage residual approach that removes level-of-development effects.

Rigidity

Online Annex Figure 1.4.7 demonstrates that countries with better public investment management practices and civil liberties, demonstrate lower spending rigidity, as better investment management systems free governments to reallocate money to where it adds most value, so budgets are less locked into past allocations. This link survives when controlled for GDP per capita.

Online Annex Figure 1.4.7 Public Spending Rigidity



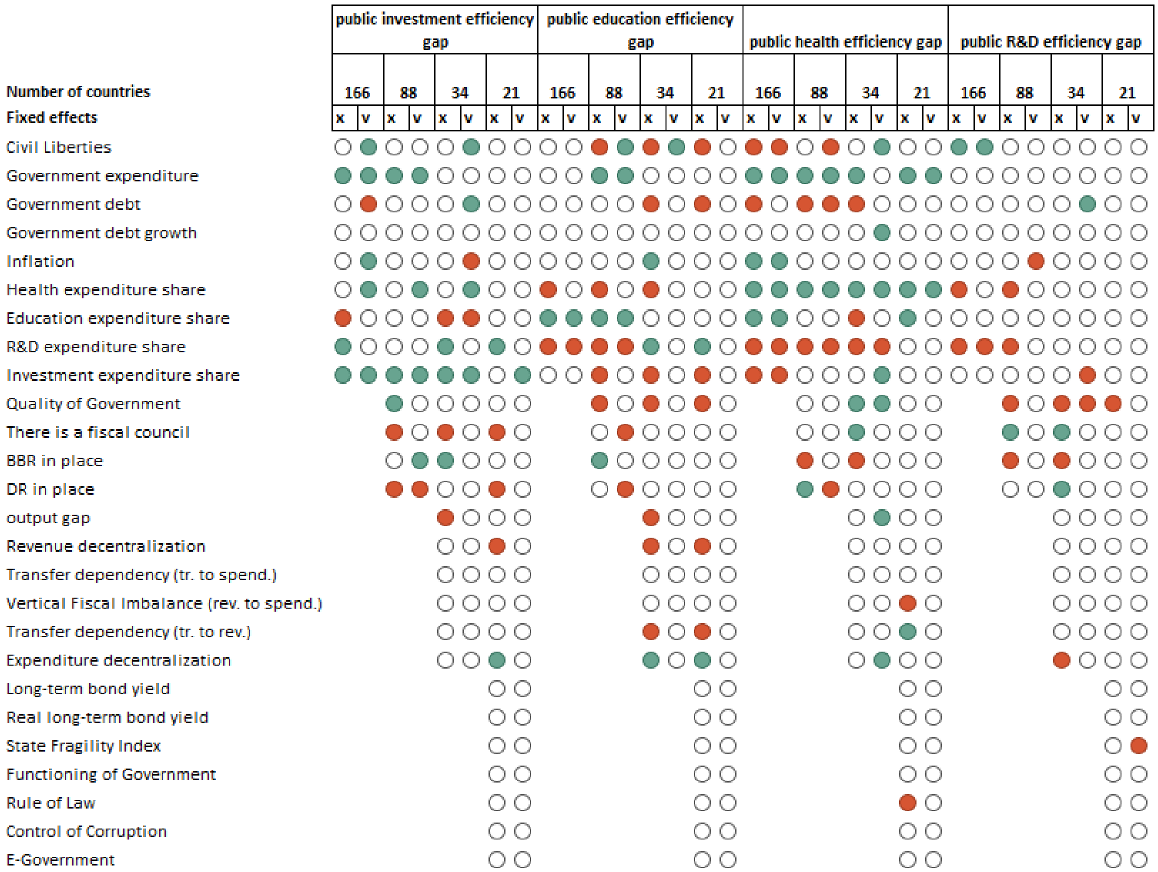


Source: Public Investment Management Assessment scores from IMF; Civil liberties from QOG dataset; Rigidity scores from IMF staff calculations.
Note: Each dot in the chart represents the average value for a given country over the sample period, providing a simplified view of the data distribution. In contrast, the fitted OLS regression line is estimated using the full panel dataset, which includes all available country-year observations. Investment management strength and civil liberties are normalized.

Robustness to Additional Controls Using BMA

The BMA results reinforce the robustness of some of the key relationships identified in the main analysis (Online Annex Table 1.4.3). For example, BMA confirms that rule of law is a robust predictor for narrower health spending efficiency gap. Decentralization shows some robustness in the case of public R&D expenditure, confirming the finding that expenditure decentralization is linked to efficiency of public spending on R&D. Finally, association of reduced public debt and higher pro-growth outlays in investment and education is confirmed by BMA as well. Other relationships do not survive this robustness check.

Online Annex Table 1.4.3 Bayesian Model Averaging Results



Source: IMF staff estimates.
Note: t-statistics are shown for each variable. Solid circles indicate coefficients with t-statistics greater than 1 in absolute value (green = positive, red = negative); hollow circles indicate insignificance. BMA is run on cross-section and panel models with and without fixed effects. There are four separate BMA analyses on four separate sets of explanatory variables; the larger the set of variables, the fewer countries have the data available. Number of countries specifies how many are left in the dataset used for the BMA, depending on the explanatory variables considered. Fixed effects “x” and “v” stands for BMA without and with fixed effects respectively.

Online Annex 1.5. Empirical Methodology for Growth Regressions[1]

This Online Annex documents the data, methodology, robustness checks, and extensions underpinning the growth regressions presented in the main text.

Data

The empirical analysis relies on a dataset of macroeconomic indicators, public expenditure components, structural reform indices, and crisis episodes. Detailed definitions, coverage, and sources are documented in *Online Annex 1.1*. The analysis also uses estimates of spending efficiency, which are detailed in *Online Annex 1.3*.

Baseline Methodology

The analysis focuses on four categories of public spending generally recognized in the literature as growth-enhancing: public investment, education, health, and research and development (R&D). To isolate compositional effects, each category is measured as a share of total public expenditure, thereby holding the overall spending envelope constant. This helps ensure that the estimated effects capture reallocation across spending categories rather than changes in the aggregate size of public spending.

The specification is event-based. Rather than regressing output on the continuous change in the spending share, the analysis constructs binary indicators for episodes of large and sustained increases in the share of each category. The use of such indicators helps identify discrete and potentially policy-driven shifts in spending composition that are less likely to be driven by automatic stabilizers or cyclical changes in the spending shares due to business cycle fluctuations.

Episodes

The analysis uses episodes of large and sustained changes in the composition of public spending which serve as the core policy shocks in both the country- and firm-level analyses. Episodes are selected to reflect durable, policy-driven reallocations in the relative composition of public spending—specifically for investment, education, health and R&D.

The baseline identification strategy defines an episode using a four-year rolling window applied to the annual change in the spending share of each category. At least three of the four years must show positive growth in the spending share, with at least one increase exceeding 1.5 standard deviations relative to the country-specific historical distribution. A single-year decline is allowed if surrounded by positive growth, to allow for slippage within longer reform periods. The statistical algorithm effectively captures relevant policy-driven shifts in spending composition. This is illustrated by the episodes of education reform in Mexico, marked by green indicators: the early 1990s saw major reforms, including the ANMEB[2] and the constitutional reform for compulsory secondary education, followed by the introduction of the PROGRESA[3] in 1997 (Online Annex Figure 1.5.1, panel 1).

The episodes indeed identify periods of sustained increases in each category of spending as a share of total spending (Online Annex Figure 1.5.1, panel 2). General public services and social protection are typically

[1] Prepared by Gabriel Hegab and Galen Sher.

[2] The *Acuerdo Nacional para la Modernización de la Educación Básica* (ANMEB) of 1992 marked a decentralization of basic education from federal to state authorities, alongside curriculum revisions and teacher incentives. In 1993, an amendment to Article 3 of Mexico's Constitution made secondary education compulsory.

[3] The *Programa de Educación, Salud y Alimentación* (PROGRESA) launched in 1997, was a pioneering conditional cash transfer program targeting poor households, linking benefits to school attendance and health checkups.

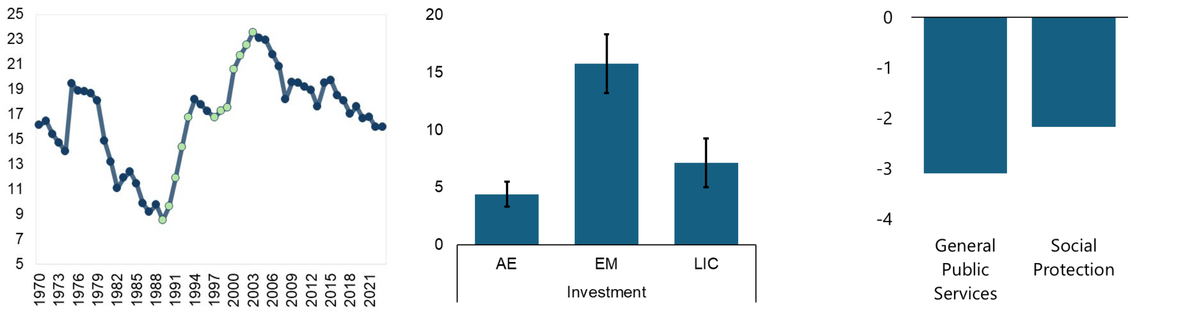
the two largest spending categories from which fiscal adjustments are drawn to accommodate increases in public investment. (Online Annex Figure 1.5.1, panel 3).

Online Annex Figure 1.5.1 Public Spending Episodes

1. Mexico: Public Spending on Education
(Percent of total expenditure)

2. Impact of Public Investment Episode on Public Investment Spending Share
(Percent)

3. Spending Reallocation During Public Investment Episodes
(Percent)



Source: IMF Government Finance Statistics (GFS) database; IMF staff calculations.
Note: Green dots indicate an episode of an increasing spending on education share.

Source: IMF staff calculations.
Note: Bars show 10-year cumulative impulse responses estimated via local projections. Shaded bands denote interquartile ranges across simulations.

Source: IMF staff calculations.
Note: Bars show 5-year cumulative impulse responses of each spending category as a share of total government expenditure, estimated via local projections.

Episodes identified through this method are relatively infrequent but span a wide range of countries and sectors, as summarized in Online Annex Table 1.5.1. They are detected across all income groups for each spending category. Coverage is broad for public investment, education and health, with representation from advanced and emerging market economies, and low-income developing countries. In contrast, public R&D spending episodes are concentrated in advanced economies, reflecting limited reallocation activity and data availability in lower-income settings.

Online Annex Table 1.5.1 Summary Statistics of Episodes by Spending Category

Statistic	Public Investment	Public Education	Public Health	Public R&D
Countries with at least one episode	138	121	155	59
Advance Economies	33	26	35	33
Emerging Markets	68	60	77	19
Low-Income Developing Countries	37	35	43	7
Total years covered by episodes	923	850	1422	402
Total number of episodes	185	172	254	76
Average episodes per country	1.3	1.4	1.6	1.3
Average episode length (years)	3.66	3.11	4.46	1.61
Countries with no episodes	56	73	39	135
Countries with 1 episode	95	79	78	43
Countries with 2 episodes	39	34	60	15

Countries with 3 episodes	4	7	12	1
Countries with 4 or more episodes	0	1	5	0
Average episodes increase in percentage point	4.3	1.6	0.8	0.1
Total countries in sample	194	194	194	194
Total years in sample	64	64	64	43
Sample year range	1960–2024	1960–2024	1960–2024	1981–2024

Public health spending shows the highest number of episodes, while public R&D spending episodes are more limited in coverage and duration, leading to less precise estimates in the case of R&D. On average, episodes last between 1.6 and 4.5 years, with most starting in the late 1990s, or early 2000s. Coverage varies by area: more than 75 percent of countries have at least one episode in public investment, education or health, while over 70 percent report no R&D episodes, reflecting data constraints and the rarity of large shifts in that category. The average size of the episodes—measured as the change in spending share between the three years before and after each episode—is 4.3 percentage points for public investment, 1.6 for education, 0.8 for health, and just 0.1 for R&D.

Growth Regressions

The chapter adopts a panel local projections approach (Jordà 2005) to estimate the medium- and long-run output dynamics associated with major compositional shifts in public expenditure.

The baseline regression specification follows the local projections framework; for each country i, year t, and horizon $h \in \{0,1,\ldots,10\}$, the following equation is estimated:

$$y_{i,t+h} - y_{i,t-1} = \theta_t^h + \alpha_i^h + \beta^h d_{i,t} + \gamma^h c_{i,t} + e_{i,t+h}^h \qquad \text{(A5.1)}$$

where $y_{i,t}$ denotes the logarithm of real GDP—or alternatively, capital, employment, or total factor productivity, $d_{i,t}$is a dummy variable equal to one in the first year of a spending episode and zero otherwise, and $c_{i,t}$ is a vector of control variables. The controls include two lags of the dependent variable for *h* equal to zero, two lags of the episode dummy to capture possible delayed effects of spending shifts, the one-year-ahead forecast of GDP growth to control for expected macroeconomic conditions and reduce endogeneity arising from forward-looking policy responses, structural reform indicators in other areas (product and labor markets from Alesina et al. 2005) to address potential omitted bias. Country and year fixed effects are included in all regressions. Estimation is performed via fixed effects panel regression with Driscoll–Kraay standard errors which are robust to general forms of heteroskedasticity, serial correlation, and cross-sectional dependence. The estimates of β^has a function of h are shown in Figure 5 panel 1-4 in the main text.

The analysis is further extended to allow for heterogeneity in the output response to pro-growth spending episodes based on the continuous level of spending efficiency. In this specification, the local projection model interacts the binary episode indicator with the efficiency score. Efficiency scores $E_{i,t}$ are estimated from a Stochastic Frontier Analysis (SFA) model applied to sector-specific spending data (see Online Annex 1.2). Specifically, the model estimates differential impulse response functions using the following specification:

$$y_{i,t+h} - y_{i,t-1} = \theta_t^h + \alpha_i^h + \beta^{high} E_{i,t} d_{i,t} + \beta^{low}(1 - E_{i,t}) d_{i,t} + \gamma^h c_{i,t} + e_{i,t+h}^h \qquad \text{(A5.2)}$$

where $E_{i,t} = 1 - \text{Efficiency Gap}_{it}$ denotes the time-varying efficiency score, bounded between 0 and 1, and $\text{Efficiency Gap}_{it}$ is explained in Online Annex 1.2.

This approach allows the marginal effect of spending episodes to vary continuously with efficiency, distinguishing the responses in more and less efficient contexts. Country fixed effects (α_i^h), year fixed effects (θ_t^h), and same baseline control variables ($c_{i,t}$) are included as in the main specification. Similarly, the estimates of β^{high} and $\beta^{low}$are reported in Figure 5 panel 6.

Robustness and Extensions

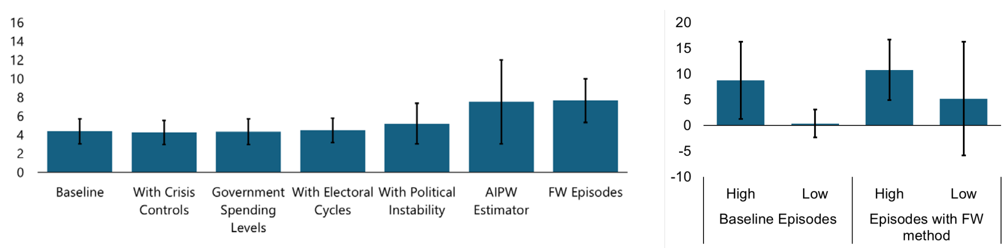
This section presents a set of robustness checks and methodological extensions. Baseline regressions are re-estimated with additional macro-fiscal controls, alternative estimation strategies, and different episode definitions.

The results obtained across alternative methods are presented in Online Annex Figure 1.5.2 and reinforce the core findings. While magnitudes differ slightly, results remain broadly consistent across specifications, suggesting that growth effects are robust across specifications and not driven by omitted variables or endogeneity. The subsection below details the robustness exercises.

Online Annex Figure 1.5.2 10-Year GDP Response to Public Investment Episodes

1. Robustness Specifications
(Percent)

2. By Low and High Efficiency Quartiles
(Percent)



Source: IMF staff calculations.
Note: Bars report cumulative GDP responses at a 10-year horizon following public investment episodes, estimated via local projections. Error bars denote 90 percent confidence intervals using Driscoll-Kraay standard errors. In the left chart, each specification adds controls or varies the estimation method: Baseline, with crisis dummies, the level of government spending, executive and legislative electoral year dummies, and a principal component index of political instability (PC) based on Bernal-Verdugo et al. (2013), augmented inverse probability weighting (AIPW), and alternative episode definition based on Forbes–Warnock (FW). In the right chart countries are grouped into high- and low-efficiency quartiles based on the average of their estimated efficiency scores.

Additional Controls

The first extension introduces controls for macroeconomic crises, including dummy variables for currency, sovereign debt, and banking crises which can trigger reforms. The spending episodes may overlap with periods of crises, potentially confounding the estimated growth effects. A second specification adds the log of total public expenditure to account for variation in the size of the fiscal envelope across countries and time. A possible concern is that compositional changes in spending may coincide with overall shifts in fiscal stance, such as deficit expansions or consolidations. Third, executive and legislative elections cycles, and political instability variables from Bernal-Verdugo et al. (2013), are included to test for robustness to potential politically-driven spending reforms. Finally, a broad set of reform shocks following Alesina et al. (2005), including trade, capital, current account, financial, product market, and labor market liberalization was also tested.

Additional Methods To Address Endogeneity

To address potential selection bias in estimating output responses to spending composition shifts, the analysis adopts the Augmented Inverse Probability Weighting (AIPW) estimator (Glyn and Quinn 2010). AIPW is a doubly robust, semi-parametric method that combines inverse probability weighting based on a propensity score model with an outcome regression model. Treatment assignment is modeled via a logit[4] regression on predetermined covariates $Z_{i,t}$, which include the spending episode and the lagged level of its spending component. These covariates are used to estimate the probability $\Pr(d_{i,t} = 1 \mid z_{i,t})$ of an episode's occurrence, generating inverse-probability weights. Higher public spending in a particular category (for example, public investment) is statistically significantly associated with a lower probability of experiencing an episode of major spending reallocation toward that category. The outcome model is estimated conditional on the treatment, where episodes are weighted by the inverse of their estimated propensity scores and a broader set of controls $c_{i,t}$ which contains (and may extend beyond) the variables in $z_{i,t}$. The AIPW estimator combines both models in a doubly robust manner, providing consistent estimates of the average treatment effect. The equations below represent the AIPW setup:

$$\Pr(d_{i,t} = 1 \mid z_{i,t}) = \Lambda(z'_{i,t}\pi), \tag{A5.3}$$

$$y_{i,t+h} - y_{i,t-1} = \theta_t^h + \alpha_i^h + \beta^h d_{i,t} + \gamma^h c_{i,t} + \varepsilon_{i,t+h}^h$$

The left panel of figure Online Annex Figure 1.5.3 shows the empirical densities of the predicted probabilities of an episode of major reallocation of public spending toward investment, for countries that experienced an episode and those that did not. The two distributions exhibit substantial overlap, suggesting that countries experiencing an episode are similar to countries that did not. Identification of causal effects through the AIPW estimator relies on this common support region, where selection into treatment most closely approximates the random allocation benchmark.

To further address potential selection bias, the chapter uses the Synthetic Control Method (SCM) of Abadie, Diamond and Hainmueller (2010). This method estimates the effect of an episode on output by comparing the evolution of output in the country undergoing the episode to that of all other countries that did not experience an episode. The counterfactual evolution of output is calculated as a weighted average of output in these other countries, using weights that match the path of output in the country of interest before the episode. The effect of the episode in year t, after the episode starts, is defined as:

$$y_{i,t} - \hat{y}_{i,t} \tag{A5.4}$$

where $y_{i,t}$ is the logarithm of GDP per person in country i that undergoes an episode and $\hat{y}_{i,t}$ is the counterfactual evolution of GDP per person in that country (i.e., how GDP per person would have evolved in the absence of the episode), defined as:

$$\hat{y}_{i,t} = \frac{1}{J-1} \sum_{j=1, j\neq i}^{J} w_j y_{j,t} \tag{A5.5}$$

where the weights w_j are non-negative, sum to 1, and minimize the distance between $y_{i,t}$ and $y_{j,t}$ in the years before the episode begins. To estimate the overall effect of a given type of reform, the chapter averages across all reform episodes, reporting the estimate

[4] A probit specification for the treatment model was also tested and found that the estimated impulse responses were nearly identical.

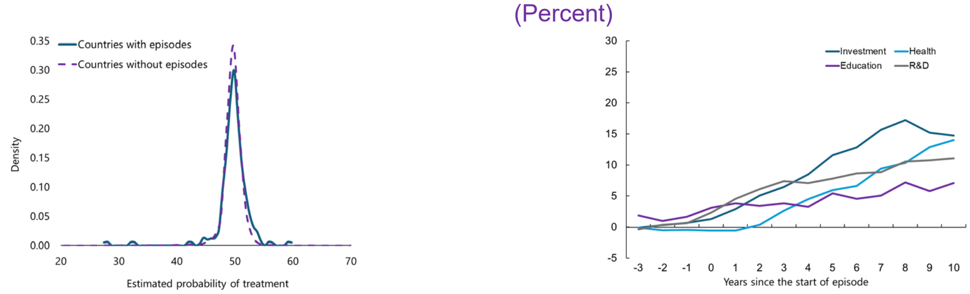
$$\frac{1}{n}\sum_{i=1}^{n}(y_{i,t} - \hat{y}_{i,t})$$

where n is the number of reform episodes.[5] The SCM confirms the large positive effect of reform episodes on output. Online Annex Figure 1.5.3 panel 2 shows that output per person is almost 15 percent higher, ten years after a major reallocation toward public investment, than in other countries that did not experience a reform episode. The effects of health, education, and R&D spending episodes are also large according to the SCM. The validity of the SCM is supported by close match between GDP per person in the treatment and control groups in the three years before the start of the episode ($t = -3, -2, -1$ along the horizontal axis).

Online Annex Figure 1.5.3 Diagnostic Tests and Results for Synthetic Control Method

1. Distribution of Estimated Treatment Probabilities
(Percent)

2. GDP Impacts of Reallocation Episodes using SCM
(Percent)



Source: IMF staff calculations.
Note: Kernel densities of estimated treatment probabilities from the first-stage logit model for episodes and non-episodes.

Source: IMF staff calculations.
Note: SCM = synthetic control method.

Alternative Episode Definitions

An alternative definition of episodes applies a 2 standard-deviation threshold (as opposed to 1.5 in the baseline) and confirms the qualitative results of the baseline. The impacts increase commensurately, which suggests that non-linearities are limited. Additionally, another alternative episode definition following Forbes and Warnock (2012) is considered, identifying periods in which annual increases exceed the five-year rolling mean by more than one standard deviation, with at least one observation surpassing two standard deviations.

Alternative Approach To Estimating Interaction Effects

The analysis also estimates heterogeneous output responses to spending episodes across discrete efficiency groups. Specifically, countries are assigned to efficiency quartiles based on the average of their estimated efficiency scores. Local projections are then re-estimated separately for each group using the following specification:

$$y_{i,t+h} - y_{i,t-1} = \theta_t^h + \alpha_i^h + \beta_q^h I\{E_i \in Q_q\} d_{i,t} + \gamma^h c_{i,t} + e_{i,t+h}^h \quad \text{(A5.6)}$$

[5] Estimation is restricted to countries that experienced one or no reform—countries with two or more reform episodes are dropped.

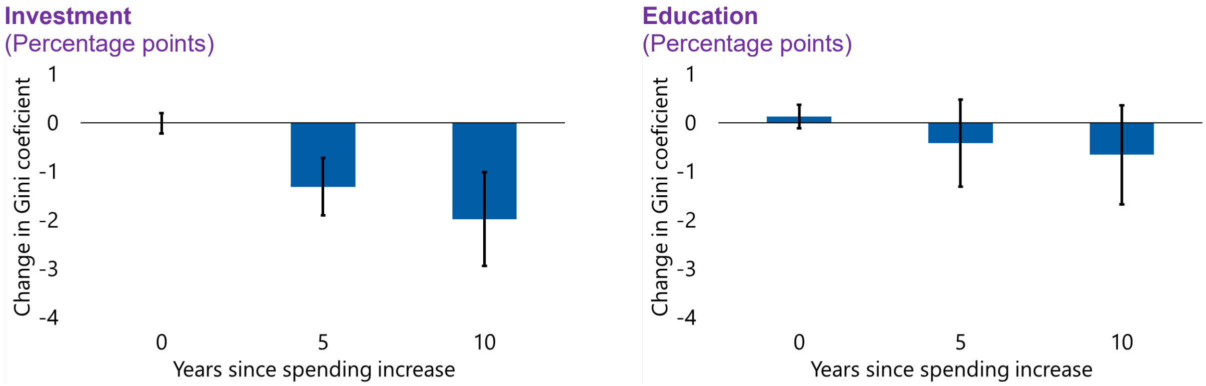
where, $E_i = T^{-1} \sum_t \exp(-\widehat{u_{i,t}})$ is the average efficiency score of country i, $q \in \{1,4\}$, and Q_q denotes quartile q of the distribution across countries of average efficiency scores. The model includes the same lagged controls and fixed effects as the baseline. This setup allows testing whether output effects are concentrated in more efficient environments or also materialize in low-efficiency groups of countries.

Impacts on Inequality

To examine the effect of spending reallocation reforms on inequality, this section estimates equation (A5.1) using the Gini coefficient of income inequality (which ranges from 0, indicating perfect equality of incomes to 100 percent, indicating maximum inequality where one individual obtains all the income) as the $y_{i,t}$ variable in place of output. The controls $c_{i,t}$ include two lags of changes in the Gini coefficient.

The results, presented in Online Annex Figure 1.5.4, indicate that the Gini coefficient tends to fall by 2 percentage points in the 10 years after a public investment reform (i.e., a major reallocation towards public investment) and by 0.7 percentage points in the 10 years after a public education spending reform. The effect of public investment is statistically significant, and the effect of education spending is marginally statistically significant, at the 10 percent level using Driscoll—Kraay standard errors. The findings for public investment extend those of Furceri and Li (2017) to episodes of increases in public investment that are financed by cuts to other types of public spending.

Online Annex Figure 1.5.4 Inequality Response to Spending Episodes



Source: IMF staff calculations.
Note: Bars report the percentage change in the Gini coefficient in response to episodes of major increases in public spending on investment and education, estimated using local projections. Error bars represent 90 percent confidence intervals computed with Driscoll–Kraay standard errors.

Implied Fiscal Multipliers

To examine the fiscal multipliers implied by the growth regressions, this section adopts the methodology of Ramey and Zubairy (2018). Extending their equation (3) to a cross-country panel data context, the chapter estimates

$$\sum_{j=0}^{h} y_{i,t+j} = \theta_t^h + \alpha_i^h + \beta^h \sum_{j=0}^{h} x_{i,t} + e_{i,t+h}^h \quad \text{(A5.7)}$$

where $y_{i,t}$ is the ratio of GDP to trend GDP and $x_{i,t}$ is the ratio of government spending in a specific category (e.g., public investment or health spending) to trend GDP. Trend GDP, in turn, is estimated for each country separately as the exponential of a 6th order polynomial in the natural logarithm of that country's GDP. Equation (A5.7) is estimated by two-stage least squares, where $\sum_{j=0}^{h} x_{i,t}$ is instrumented by $d_{i,t}$, the dummy variable described above that is equal to one in the first year of a spending episode and zero otherwise. The estimate of β^h is an estimate of the cumulative multiplier from this specific spending category at horizon h (i.e., h years after the start of the reform episode). These estimated multipliers are reported in Footnote 18 of the main text.

Online Annex 1.6. Theoretical Model of Public Spending[1]

The appendix presents the modelling framework used to analyze the long-term impacts of closing public spending efficiency gaps and reallocating public spending toward more pro-growth areas.

The model's novelties lie in its integration of expenditure-side fiscal policy components and human capital development within an endogenous growth framework. It brings together several strands of literature:

- *Public infrastructure investment.* The framework builds on the literature on public infrastructure investment and output multipliers, as in Traum and Yang (2015), Drautzburg and Uhlig (2015), and IMF (2020). Government investment adds to the stock of productive public capital—such as roads or communication networks—which, in turn, enhances private sector productivity.
- *Human capital development.* Building on Blankenau and Simpson (2004), Dissou et al. (2016), and Atolia et al. (2017), private human capital is modeled as the outcome of households' allocation of time to education and healthcare. This process is complemented by public investment in human capital, which adds to the stock of public human capital—such as schools and hospitals—that enhances the effectiveness of private efforts.
- *Endogenous growth.* Following Comin and Gertler (2006) and Anzoategui et al. (2019), the model distinguishes between the stock of created technologies, Z_t, that is available for adoption, and the stock of adopted technologies, A_t. Public R&D expenditures contribute to the creation of new technologies by increasing Z_t, while private adoption efforts expand A_t. This distinction between creation and adoption allows for realistic delays in technology diffusion, and thus more accurately reflects the dynamic impact of public R&D spending. The gap $Z_t - A_t$ represents the stock of unadopted technologies.
- *Efficiency-adjusted public expenditure.* The accumulation of both infrastructure and human capital stocks is adjusted for inefficiencies, as in Berg et al. (2019). This implies that only a fraction of public investment is transformed into productive capital. A similar adjustment is applied to public R&D spending, reflecting inefficiencies in translating expenditure into technological advancement.

Modelling framework

Overall setup

The model is based on a New Keynesian framework, extended to explicitly incorporate human capital accumulation and endogenous growth through technology adoption. The economy is populated by a representative household, a government, and a multi-tiered production sector. The representative household makes optimal decisions regarding consumption, labor supply, private physical investment, and savings in government bonds. Departing from the standard framework, the household also invests in human capital (health and education). The government comprises a fiscal authority and a central bank. The fiscal authority's role includes expenditures on public infrastructure, public human capital, and public R&D, in addition to standard government consumption, lump-sum transfers, and taxes levied on consumption and income. The central bank sets the nominal interest rate according to a Taylor rule.

The model's production structure follows Comin and Gertler (2006) and Anzoategui et al. (2019), comprising three tiers:

[1] Prepared by Krzysztof Bańkowski and Anh Dinh Minh Nguyen.

- *Final goods.* A monopolistically competitive sector produces differentiated goods and sets nominal prices according to a Calvo-style staggered pricing mechanism.
- *Intermediate goods.* This sector consists of a continuum of firms, each operating Cobb-Douglas production functions using effective labor (defined as the product of private human capital and hours worked) and private capital, with public infrastructure capital entering as a positive externality.
- *Technology adopters.* A competitive set of "adopters" converts unadopted technologies $Z_t - A_t$ for use in production. The probability of successful adoption in a given period depends positively on the adopter's level of adoption expenditure, measured in units of the final consumption goods composite. Adoption success is modeled as a probabilistic process: if unsuccessful, the adopter may attempt conversion again in subsequent periods. Once a technology is successfully adopted, the adopter sells the rights to use technology to intermediate-goods producers.

Key equations

Production. The aggregate production function in the economy is given by:

$$Y_t = A_t^{\vartheta-1} K_t^{\alpha} (H_t L_t)^{1-\alpha} (K_t^{GI})^{\alpha_G}, \tag{Eq.6.1}$$

where Y_t denotes output, A_t is the stock of adopted technologies, K_t is the private capital stock, $H_t L_t$ is effective labor—which is the product of private human capital H_t and hours worked L_t, and K_t^{GI} is the stock of public infrastructure capital. The parameter ϑ governs the productivity impact of technology adoption, while α and α_G denote the output elasticities of private capital and public infrastructure, respectively.[2]

Private human capital accumulation. Households divide their unit time endowment among time spent on human capital accumulation (E_t), labor supply (L_t), and leisure ($1 - E_t - L_t$). The evolution of private human capital H_t is described by the following law of motion:

$$\mathrm{H}_{t+1} = \left(1 - \delta^{\mathrm{h}}\right) \mathrm{H}_t + \lambda(\mathrm{E}_t)^{\gamma} \left(\frac{K_t^{GE}}{\Lambda_t}\right)^{\mu}, \tag{Eq.6.2}$$

where δ^h is the depreciation rate of human capital, E_t denotes the time spent by the household in human capital development, and K_t^{GE} represents the stock of public human capital (e.g., schools and hospitals). The term Λ_t is a scaling factor ensuring consistency with balance growth. The cost of time spent on working and human capital development is a reduction in leisure, which lowers utility.

Public capital accumulation and efficiency. The accumulation of public infrastructure capital (K_t^{GI}) and public human capital (K_t^{GE}), adjusted for corresponding inefficiencies e^{GI} and e^{GE}, is expressed as follows:

$$K_t^{GI} = (1 - \delta^{GI})\, K_{t-1}^{GI} + (1 - e^{GI})\, I_t^{GI}, \tag{Eq.6.3}$$

$$K_t^{GE} = (1 - \delta^{GE})\, K_{t-1}^{GE} + (1 - e^{GE})\, I_t^{GE}. \tag{Eq.6.4}$$

[2] As shown in Anzoategui et al. (2019), ϑ reflects the intermediate-goods elasticity of substitution.

The inefficiency parameters e^{GI} and e^{GE} capture the concept of technical efficiency—that is, only a fraction 1-e^{x} of each unit of spent contributes to the respective capital stock, in a similar way with the efficiency gap discussed in Online Annex 1.2.[3]

Technological development. The model distinguishes between the stock of available technologies Z_t and the stock of adopted technologies A_t. The evolution of Z_t depends on two main factors. First, higher levels of human capital are assumed to foster technological innovation—that is, increase in skilled labor leads to more ideas and innovations, reflecting a learning-by-doing channel as discussed in Chang et al. (2002). Second, public R&D spending (G_t^{RD}) contributes directly to the creation of new technologies.

$$log\left(\frac{Z_t}{Z^{SS}}\right) = \rho_A\ log\left(\frac{Z_{t-1}}{Z^{SS}}\right) + (1-\rho_A)\left(\alpha_{HA}\ log\left(\frac{H_{t-1}}{H^{SS}}\right) + \alpha_{RD}\ log\left(\frac{G_{t-1}^{RD}}{G^{RD,SS}}\right)\right) \quad \text{(Eq.6.5)}$$

To capture the lags between creation and adoption, the evolution of adopted technologies A_t is modeled as:

$$A_{t+1} = \lambda_t\phi[Z_t - A_t] + \phi A_t \quad \text{(Eq.6.6)}$$

where the time-varying probability of adoption is given by $\lambda_t = \lambda_0\left(\frac{A_tS_t}{\Lambda_t}\right)^{\varsigma}$ and depends positively on adopters' expenditure S_t, scaled by the balanced-growth trend Λ_t. Adopters choose S_t to maximize the value of converting unadopted technologies into usable ones, taking into account the associated cost.

Model Parameterization

The model is parameterized for a representative advanced economy (AE) and a representative emerging market and developing economy (EMDE), separately. Key parameters are briefly described below and summarized in Online Annex Table 1.6.1.

- *Elasticity of output with respect to infrastructure* (α_G). It is set at 0.05 for AEs, following Traum and Yang (2015), and 0.17 for EMDEs, based on Born and Ligthart (2019) and Buffie et al. (2012). The higher value for EMDEs reflects the larger contribution of public infrastructure investment to output in these economies.
- *Private human capital accumulation.* The elasticity of human capital accumulation with respect to the public human capital stock (μ) is set at 0.25 for EMDEs , and 0.1 for AEs, which is in the range documented in Glomm and Ravikumar (1997), Blankeau (2004), and Atolia et al. (2016). The lower value for the latter reflects the empirical finding that the marginal impact of government investment on human capital formation is moderate, given already well-developed education and health infrastructure. The elasticity with respect to the time spent on human capital accumulation is set to $\gamma = 0.5$ for both country groups, consistent with Atolia et al. (2016) and Dissou et al. (2016).

Online Annex Table 1.6.1 Selected Model Parameters

Description	AE	EMDE
Parameters		
α_G	0.05	0.17
α_{HA}	0.45	-
α_{RD}	0.1	-
ρ_A	0.79	0.79
ϕ	0.97	0.97
e^{GI}	0.35	0.42
e^{GE}	0.30	0.32
e^{GRD}	0.41	-
ς	0.8	0.8
λ_0	0.1	0.1
δ^{GI}, δ^{GE}	0.025	0.025
μ	0.1	0.25
γ	0.5	0.5

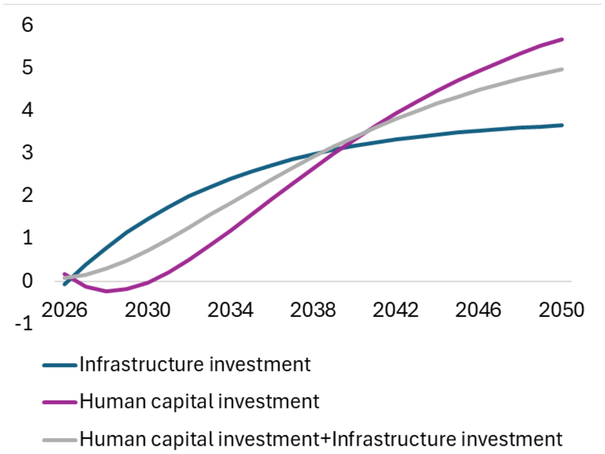
[3] Another dimension of efficiency relates to the allocation of resources toward higher-return investment projects– referred to as allocative efficiency. The following section also considers an extended scenario that incorporates both technical and allocative efficiency improvements.

- *Inefficiency parameters*. The efficiency gaps are set based on empirical estimates in the chapter, as explained in the main text. For infrastructure investment, the average gap is set at 0.35 for AEs and 0.42 for EMDEs. For human capital investment, they are 0.30 and 0.32 for AEs and EMDEs, respectively. For public R&D spending, the inefficiency parameter is set at 0.4 for AEs, while this innovation channel is shut down in EMDEs due to their reliance on technology adoption rather than on innovation.
- *Depreciation of public capital* (δ^{GI}, δ^{GE}). Both depreciation rates are set at the standard value of 0.025, corresponding to an annual depreciation rate of 10 percent.
- *Technology innovation parameters*. The AR(1) coefficient for the stock of technologies (ρ_A) is set at 0.79 (quarterly), as in Chang et al. (2002). For AEs, the elasticity of human capital with respect to technology development (α_{HA}) is set at 0.45, capturing the learning-by-doing channel, as in Chang et al. (2002). The elasticity of public R&D spending (α_{RD}) is calibrated at 0.1, consistent with the R&D impact estimates in Fieldhouse and Mertens (2025).
- *Technology adoption parameters*. Following Comin and Gertler (2006) and Anzoategui et al. (2019), the elasticity of the probability of adoption with respect to adoption spending (ς) is set at 0.8, and λ_0 is set at 0.1, implying an average lag of approximately ten years.

In addition, the steady states of fiscal variables are set as follows:

- *Public consumption-to-GDP ratio*. Set to standard values for each country group, 18 percent of GDP for AEs and 14 percent of GDP for EMDEs, based on IMF (2023) and based on World Bank's World Development Indicators database. Public consumption includes wages of teachers, doctors and nurses, other public workers employed in education and health care, as well as other government final consumption.
- *Public investment-to-GDP ratio*. Infrastructure investment is set at 3 percent of GDP for AEs and 5 percent of GDP for EMDEs (excluding public education and health investment, like building schools and hospitals). Public human capital spending (investing in public education and health) is calibrated at 1.5 percent of GDP for AE and 2 percent of GDP for EMDEs, based on the share of these items in public investment of OECD income group (2025a) due to limited data availability of public investment composition.
- *Effective tax rates*. Effective tax rates are set at standard levels for each country group, following IMF (2023), Benitez et al. (2023), and (OECD, 2021, 2025b). The consumption tax rate is set at 18 percent for AEs and 15 percent for EMDEs. The income tax rate is 25 percent for AEs and 10 percent for EMDEs.

Online Annex Figure 1.6.1 Emerging Market and Developing Economies: Output Gain from Combining Human Capital and Infrastructure
(Percent deviation from steady state)



Source: IMF staff estimates.
Note: Responses of output to a permanent increase of 1 percent of GDP in 2025 in public human capital investment, public infrastructure investment, and a 50/50 mix of both, financed by an equivalent cut to public consumption.

Additional Results

Implied Multipliers

The implied cumulative multipliers over 25 years for public investment for human capital investment and infrastructure investment are about 1 for AEs and 2.5 for EMDEs, which are in the range of long-run multipliers by Ramey (2020) (Table 3). Afonso and Eduardo (2024) also document that the multiplier of public investment is 0.9 for advanced economies and 2.3 for emerging economies. Meanwhile, the cumulative multiplier for R&D expenditure is about 1.8, as in line with Fieldhouse and Mertens (2025).

Policy-Mix for Emerging Market and Developing Economies

As shown in the main text, reorienting resources toward human capital investment and infrastructure investment can result in long term gain, though over different time horizons. A combination of human capital investment and infrastructure investment could be also a viable option for emerging markets and developing economies. This policy mix results in a more balanced outcome by combining the short-term gains from infrastructure investment and longer-term gains from human capital investment. While this option delivers a slightly lower output gain by 2050 compared to investing more solely in human capital, it yields higher output in the shorter term.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490): 493–505.

Afonso, António, and Eduardo Rodrigues. 2024. "Is Public Investment in Construction and in R&D, Growth Enhancing? A PVAR approach." *Applied Economics* 56.24: 2875–99.

Afonso, António, João Tovar Jalles, and Ana Venâncio. 2023. *Government Spending Efficiency, Measurement and Applications: A Cross-Country Efficiency Dataset*: 44–71. Cheltenham, UK: Edward Elgar Publishing.

Aigner, Dennis, C. A. Knox Lovell, and Peter Schmidt. 1977. "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics* 6(1): 21–37.

Alesina, A., S. Ardagna, G. Nicoletti, and F. Schiantarelli. 2005. "Regulation and Investment." *Journal of the European Economic Association*, 3(4): 791–825.

Alesina, Alberto, Davide Furceri, Jonathan D Ostry, Chris Papageorgiou, and Dennis P Quinn. 2024. "Structural Reforms and Elections: Evidence from a World-Wide New Dataset." *Journal of the European Economic Association* 22(4) 1936–80.

Anzoategui, Diego, Diego Comin, Mark Gertler, and Jose Martinez. 2019. "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence." *American Economic Journal: Macroeconomics* 11(4): 67–110

Apeti, Augustine E., Benedict Bambe, and Ali A. B. Lompo. 2023. "Determinants of Public Sector Efficiency: A Panel Database from a Stochastic Frontier Analysis." *Oxford Economic Papers* 76(3): 741–58.

Atolia, Manoj, Benjamin G. Li, Ricardo Marto, and Giovanni Melina. 2021. "Investing in Public Infrastructure: Roads or Schools." *Macroeconomic Dynamics* 25(7): 1827–63.

Barro, Robert J, and Jong-Wha Lee. 1993. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics, 104:* 184–98.

Baum, Anja, Tewodaj Mogues, and Geneviève Verdier. 2020. "Getting the Most from Public Investment." Chapter 3 in *Well Spent: How Strong Infrastructure Governance Can End Waste in Public Investment*, edited by Gerd Schwartz, Manal Fouad, Torben Hansen, and Geneviève Verdier. Washington, DC: International Monetary Fund.

Benitez, Juan Carlos, Mario Mansour, Miguel Pecho, and Charles Vellutini. Building Tax Capacity in Developing Countries. Staff Discussion Note, International Monetary Fund, 2023.

Berg, Andrew, Edward F. Buffie, Catherine Pattillo, Rafael Portillo, Andrea F. Presbitero, and Luis-Felipe Zanna. 2019. "Some Misconceptions about Public Investment Efficiency and Growth." *Economica* 86(344): 697–24.

Bernal-Verdugo, Lorenzo E., Davide Furceri, and Dominique Guillaume. 2013. "The Dynamic Effect of Social and Political Instability on Output: The Role of Reforms." IMF Working Paper 13/91, International Monetary Fund, Washington, DC.

Blankenau, William, 2005. "Public Schooling, College Subsidies and Growth." *Journal of Economic Dynamics and Control*, 29(3):487–507.

Blankenau, William, and Nicole Simpson. 2004. "Public Education Expenditures and Growth." *Journal of Development Economics* 73 (2): 583–605.

Born, Benjamin, and Jenny Ligthart. 2014. "Public Infrastructure Investment and Economic Performance: A Meta-Analysis." *Journal of Economic Surveys* 28 (5): 799–819.

Buffie, Edward, Andrew Berg, Catherine Pattillo, Rafael Portillo, and Luis-Felipe Zanna. 2012. "Public Investment, Growth, and Debt Sustainability: Putting Together the Pieces." IMF Working Paper No. 12/199, International Monetary Fund, Washington, DC.

Chang, Yongsung, Joao F. Gomes, and Frank Schorfheide. 2002. "Learning-By-Doing as a Propagation Mechanism." *American Economic Review* 92(5): 1610–37.

Comin, Diego, and Mark Gertler. 2006. "Medium-Term Business Cycles." *American Economic Review* 96 (3): 523–51.

Dissou, Yazid, Selma Didic, and Tatsiana Yakautsava. 2016. "Government Spending on Education, Human Capital Accumulation, and Growth." *Economic Modelling* 58: 499–511.

Drautzburg, Timo, and Harald Uhlig. 2015. "Fiscal Stimulus and Distortionary Taxation." *Review of Economic Dynamics* 18(4): 860–80.

Fieldhouse, Andrew, and Karel Mertens. 2023. "The Impact of Public R&D Spending on Innovation." NBER Working Paper 30894, National Bureau of Economic Research, Cambridge, MA.

Forbes, K. J., and F. E. Warnock. 2012. "Capital Flow Waves: Surges, Stops, Flight, and Retrenchment." *Journal of International Economics*, 88(2): 235–51.

Furceri, Davide, and Bin Grace Li. 2017. "The Macroeconomic (and Distributional) Effects of Public Investment in Developing Economies." IMF Working Paper WP/17/217, International Monetary Fund, Washington, DC.

Garcia-Escribano, Mercedes, Tewodaj Mogues, and Pablo Juarros. 2022. "Patterns and Drivers of Health Spending Efficiency." IMF Working Paper 2022/048, International Monetary Fund, Washington, DC.

Gethin, Amory. 2024. "A New Database of General Government Revenue and Expenditure by Function, 1980–2022."

Glomm, Gerhard, and Brinda Ravikumar. 1998. "Flat-Rate Taxes, Government Spending on Education, and Growth." *Review of Economic Dynamics* 1(1): 306–25.

Glyn, A., and K. Quinn. 2010. "An Introduction to the Augmented Inverse Propensity Weighted Estimator." *Political Analysis*.

Greene, William. 2005. "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model." *Journal of Econometrics* 126(2): 269–303.

Herrera, Santiago, and Eduardo Olaberria. 2020. "Budget Rigidity in Latin America and the Caribbean—Causes, Consequences, and Policy Implications." International Development in Focus. Washington, DC: World Bank.

Herrera, Sergio, Haruki Isaka, and Aristide Ouedraogo. 2025. "Efficiency of Public Spending in Education, Health, and Infrastructure: An International Benchmarking Exercise." *Journal of Applied Economics* 28 (1).

International Monetary Fund (IMF). 2016. "Managing Government Compensation and Employment - Institutions, Policies, and Reform Challenges." Policy Paper, International Monetary Fund, Washington, DC:

International Monetary Fund (IMF). 2020. "Fiscal Monitor: Policies for the Recovery." International Monetary Fund, Washington, DC:

International Monetary Fund (IMF). 2021. "IMF Investment and Capital Stock Dataset. (ICSD) 2021: Manual & FAQ - Estimating public, private, and PPP capital stocks." International Monetary Fund, Washington, DC:

International Monetary Fund (IMF). 2023. "Fiscal Monitor: Climate Crossroads: Fiscal Policies in a Warming World." W International Monetary Fund, Washington, DC:

Jaramillo, Laura, Aliona Cebotari, Yoro Diallo, Rhea Gupta, Yugo Koshima, Chandana Kularatne, Jeong Dae Lee, Sidra Rehman, Kalin I. Tintchev, and Fang Yang. 2023. "Climate Challenges in Fragile and Conflict-Affected States." Staff Climate Note 2023/001, International Monetary Fund, Washington, DC.

Jordà, Ò. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95(1): 161–82.

Kapsoli, Josefina, and Tewodaj Mogues. 2023. "Benchmarking Infrastructure Using Public Investment Efficiency Frontiers." IMF Working Paper 2023/101, International Monetary Fund, Washington, DC.

Laeven, Luc, and Fabian Valencia. 2018. "Systemic Banking Crises Revisited." IMF Working Paper WP/18/206, International Monetary Fund, Washington, D.C.

MacKinnon, James G., and Hilbert White. 1985. "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics* 29: 305–25.

Mattina, D. Todd, and Victoria Gunnarsson. 2007. "Budget Rigidity and Expenditure Efficiency in Slovenia." IMF Working Paper 2007/131, International Monetary Fund, Washington, D.C.

OECD. 2021. "Taxing Wages 2019-2020. Taxing Wages in Selected Partner Economies: Brazil, China, India, Indonesia and South Africa in 2019." OECD Publishing, Paris.

OECD. 2025a. "Government at a Glance 2025." Table J.10.4. Structure of general government investment by function. OECD Publishing, Paris.

OECD. 2025b. "Taxing Wages 2025: Decomposition of Personal Income Taxes." OECD Publishing, Paris.

Piguillem, Facundo, and Alessandro Riboni. 2024. "Sticky Spending, Sequestration, and Government Debt." *American Economic Review* 114(11): 3513–550.

Ramey, Valerie A. 2020. "The Macroeconomic Consequences of Infrastructure Investment." NBER Working Paper 27625, Cambridge, MA: National Bureau of Economic Research.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data." *Journal of Political Economy* 126(2): 850–901.

Steel, Mark F J. 2020. "Model Averaging and Its Use in Economics." *Journal of Economic Literature* 58(3): 644–719.

Traum, Nora, and Shu-Chun S. Yang. 2015. "When Does Government Debt Crowd Out Investment?" *Journal of Applied Econometrics* 30(5): 783–804.