FAMILY PLANNING-SUSTAINABLE DEVELOPMENT GOALS MODEL

MODELING THE EFFECTS OF FAMILY PLANNING ON THE SUSTAINABLE DEVELOPMENT GOALS

Family Planning-Sustainable Development Goals Model Methodology and User’s Manual
Acknowledgments

The development of the Family Planning-Sustainable Development Goals (FP-SDGs) Model benefited from assistance and guidance from a number of people and institutions. The U.S. Agency for International Development (USAID) funded the research, design, validation, and initial applications of the FP-SDGs Model through the Health Policy Plus (HP+) project—the support of Linda Cahaelen, Margaret Reeves, and Kelly Thomas was invaluable. With USAID, several subject matter experts in demography and economics offered key technical feedback and suggestions during the model development process. These individuals included John Bongaarts, John Stover, Elizabeth Leahy Madsen, Jason Bremner, Katharine Kripke, Jacob Adetunji, Bill Winfrey, and Stan Bernstein. Finally, several Palladium colleagues contributed significantly throughout the model creation process, including Jay Gribble, Bernice Kuang, Matthew Deas, Sayaka Koseki, Kathryn Corpuz, and Sara Pappa.
Abbreviations

ASFR    age-specific fertility rate
Cc       contraception index
CFI      comparative fit index
Ci       insusceptibility index
Cm       marriage index
CPR      contraceptive prevalence rate (modern and traditional)
FP       family planning
GCI      Global Competitiveness Index
GDP      gross domestic product
HP+      Health Policy Plus
ICT      information and communication technology
Ln       natural log
mCPR     modern contraceptive prevalence rate
MDG      Millennium Development Goal
MMR      maternal mortality ratio
N        sample size
RMSEA    root mean squared error of approximation
SDG      Sustainable Development Goal
SEM      structural equation modeling
Sqrt     square root
t        time (baseline)
t+1      time (each subsequent year in model)
TF       total fecundity
TFP      total factor productivity
TFR      total fertility rate
U5MR     under-five mortality rate
USAID    U.S. Agency for International Development
USD      U.S. dollars
Executive Summary

On September 25, 2015, member states of the United Nations adopted the Sustainable Development Goals (SDGs), a new global agenda to end poverty, protect the planet, and ensure prosperity for all by 2030. While the SDGs are not legally binding, governments are expected to take ownership and establish national frameworks to achieve the 17 goals—a challenging prospect given competing government financial and programmatic priorities.

In response, the Health Policy Plus (HP+) project, funded by the U.S. Agency for International Development, developed a projection tool that enables users to simulate the country-level effects of increasing family planning use—considered one of the most cost-effective SDG targets (Copenhagen Consensus Center, 2016)—on 13 health and non-health SDG indicators by 2030 and 2050. With the Family Planning-Sustainable Development Goals (FP-SDGs) Model, users can quantify the boost family planning offers toward meeting the SDGs, based on different levels of program effort, thus enabling more women, adolescents, and couples to use contraception.

User-defined base year data inputs combine with three user-created future scenarios for the country of interest, generating population projections for each model year. These population projections interact with statistically derived equations to quantify the boost family planning offers for the 13 SDG indicators. Projected outcomes relate to poverty, food security, child stunting, educational achievement, water and sanitation services, income growth, child labor, and others.

This technical guide describes the rationale and design of the FP-SDGs Model, including a step-by-step, detailed manual that explains how to apply the model. The final section details the evidence justifying the selection of explanatory variables, as well as the equations for each relationship used in the model. The guide concludes with an annex that has information on the definitions and sources for all the variables.
Introduction

On September 25, 2015, 193 member-states of the United Nations adopted the Sustainable Development Goals (SDGs), a new global agenda to end poverty, protect the planet, and ensure prosperity for all by 2030. The SDGs are an ambitious set of 17 goals, 169 associated targets, and 232 indicators pursued through national action and international cooperation (United Nations, 2015). In response to the challenge of prioritizing and implementing such a large development agenda, the Post-2015 Consensus assessed the costs and benefits of the SDG targets. It identified universal access to family planning as one of the smartest SDG targets, with large social, economic, and environmental benefits per dollar spent (Copenhagen Consensus Center, 2016). Other studies have similarly identified family planning programs as a highly cost-effective intervention (Horton and Levin, 2016; Singh et al., 2014).

In the context of competing financial and programmatic priorities and a changing international aid landscape, the Health Policy Plus (HP+) project, funded by the U.S. Agency for International Development (USAID), developed a model that projects the country-level effects of contraceptive use—using demographic change—on various SDG indicators. The resulting Family Planning-Sustainable Development Goals (FP-SDGs) Model enables users to quantify the boost family planning offers toward realizing the SDGs, given different levels of program effort; enabling more women, adolescents, and couples to use contraception. To accurately describe and project outcomes, the model is not limited to family planning or population inputs alone, but addresses the interplay between them, as affected by programmatic and policy variables across other sectors. In addition to being comprehensive in scope, the model is based on a foundation of empirical and statistical research.

Users create three scenarios for their country of interest, setting future values for family planning, education, governance, economic growth, and other policy variables. Resulting population projections interact with equations to quantify the boost family planning offers toward realizing individual SDG indicators for each model year. The model allows a comparison of scenarios to show the additive benefits of different combinations of investments. The 13 projected outcomes relate to poverty, food security, child stunting, educational achievement, water and sanitation services, income growth, child labor, and others (see Figure 1).

The FP-SDGs Model builds on the theory and statistical relationships from DemDiv, a continually applied model developed under the USAID-funded Health Policy Project. A cross-national, customizable projection model, DemDiv can be used to inform policymakers in high-fertility countries about the potential benefits of the demographic dividend, generating support for investments in the multisectoral policies required to achieve those benefits. The idea for the FP-SDGs Model came from the success and utility of Millennium Development Goals (MDGs) analyses. Using a cost-benefit analysis approach, results showed that family planning was a strong complement to—rather than a tradeoff with—other health, development, and poverty reduction efforts embedded in the MDGs.

With DemDiv and the MDGs analyses, the FP-SDGs Model joins a variety of HP+ and Health Policy Project tools intended to help advocates, policymakers, donors, and program staff determine the cost and impact of investing in family planning. For a detailed description of these previous models, see the Crosswalk of Family Planning Tools.
Figure 1. Sustainable Development Goals Indicators Included in the FP-SDGs Model Outputs

<table>
<thead>
<tr>
<th>No.</th>
<th>Indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.1</td>
<td>Proportion of the population below the international poverty line</td>
<td></td>
</tr>
<tr>
<td>2.1.2</td>
<td>Prevalence of moderate or severe food insecurity</td>
<td></td>
</tr>
<tr>
<td>2.2.1</td>
<td>Prevalence of stunting among children under 5 years of age</td>
<td></td>
</tr>
<tr>
<td>3.1.1</td>
<td>Maternal mortality ratio</td>
<td></td>
</tr>
<tr>
<td>3.2.1</td>
<td>Under-5 mortality rate</td>
<td></td>
</tr>
<tr>
<td>3.7.2</td>
<td>Adolescent birth rate per 1,000 women in that age group</td>
<td></td>
</tr>
<tr>
<td>4.1.1</td>
<td>Proportion of children at the end of primary school achieving at least a minimum proficiency level in reading</td>
<td></td>
</tr>
<tr>
<td>6.1.1</td>
<td>Proportion of the population using safely managed drinking water services</td>
<td></td>
</tr>
<tr>
<td>6.2.1</td>
<td>Proportion of the population using safely managed sanitation services</td>
<td></td>
</tr>
<tr>
<td>8.1.1</td>
<td>Annual growth rate of real GDP per capita</td>
<td></td>
</tr>
<tr>
<td>8.2.1</td>
<td>Annual growth rate of real GDP per employed person</td>
<td></td>
</tr>
<tr>
<td>8.7.1</td>
<td>Proportion of children aged 5–17 years engaged in child labor</td>
<td></td>
</tr>
<tr>
<td>11.1.1</td>
<td>Proportion of urban population living in slums, informal settlements, or inadequate housing</td>
<td></td>
</tr>
</tbody>
</table>


Objectives of the FP-SDGs Model

The FP-SDGs Model is an evidence-based advocacy tool that offers support in family planning decision-making at the country-level. Applying the model in-country answers two key questions: (1) why does family planning matter for the SDGs, and (2) to what extent can increasing family planning use help achieve the SDGs? By showcasing the multisectoral benefits of contraceptive use, the model can generate broad-based support for the following:

- Increasing financial investments in family planning programs overall, or specific programmatic line items, including commodities and consumables
- Revising or updating family planning policies and strategies, as well integrating family planning components in the policies and strategies of other sectors (e.g., including behavior change communication activities and other community-level initiatives related to family planning, population, and development)
- Bolstering policy and program implementation
- Building political will to support family planning
Model Structure

The FP-SDGs Model has four core modules: (1) demographics, (2) health, (3) economics, and (4) development. In each module, base year data and future values for scenarios serve as equation inputs, enabling the user to derive values for each SDG outcome indicator for all model projection years. The definitions of variables requiring user inputs are included in the model and in Annex A. The following is an overview of the modules and their diagrammatic schemes (see the Methodology section for detailed information about each module).

The demographics module (see Figure 2) is the core of the FP-SDGs Model. It draws on the base year inputs in family planning and other proximate determinants of fertility (e.g., marriage/union and postpartum insusceptibility), as well as male and female life expectancy to derive the annual future total fertility and life expectancy. These data then feed into the DemProj module of Spectrum to produce population projections for each of the user-defined model years. Demographic outcomes include total population; population across various age groups (e.g., ages 1–4, 5–14, 15–64, 15+), annual number of births, adolescent birth rate, total labor force, and child dependency ratio.

In the model, the proportion of demand for family planning satisfied with modern methods is calculated based on the user-entered contraceptive prevalence rate (CPR) and unmet need information. The adolescent birth rate is derived from the total fertility rate (TFR).

Figure 2. Demographics Module

KEY
User Input; Calculated; SDG indicator

$t = \text{time (base year)}$; $t+1$ represents each model year beyond the base year

---

1 Set of biological and behavioral factors through which social, economic, and environmental variables affect/inhibit fertility.
The *health module* (see Figure 3) contains computations for four SDG indicators: prevalence of moderate or severe food insecurity, prevalence of stunting, maternal mortality ratio (MMR), and under-five mortality rate (U5MR). Drawing on user base year inputs, future scenario values, as well as outputs from the demographics module, these outcomes are derived using multiple regression and structural equation modeling (as outlined in the Methodology section).

**Figure 3. Health Module**

![Health Module Diagram]

**KEY**

*User Input; Calculated; SDG Indicator*

\(t = \text{time (base year)}; \ t+1 \text{ represents each model year beyond the base year} \)
The computations from the *economics module* (see Figure 4) are largely replicated from DemDiv. Specifically, gross domestic product (GDP) is calculated as the outcome of three sub-equations: capital stock, human-capital-augmented employment, and total factor productivity (TFP). Each component is estimated through multiple regression, using base year data and outputs from the demographics module. One outcome in this section—the proportion of the population below the international poverty line—is estimated through multiple regression and was not included in DemDiv.

**Figure 4. Economics Module**

![Diagram of the economics module](https://example.com/diagram.png)

**KEY**

*User Input; Calculated; SDG Indicator*

\( t = \text{time (base year); } t+1 \text{ represents each model year beyond the base year} \)
Finally, the development module (see Figure 5) features five SDG outcomes, and is unique to the FP-SDGs Model. These relationships are computed through multiple regression, drawing from results from the demographics module, as well as user inputs and scenario assumptions.

Figure 5. Development Module

KEY: User Input; SDG Indicator
\( t \) = time (base year); \( t+1 \) represents each model year beyond the base year
Methodology

The equations relating family planning use and demographic indicators to health, economic, and other development outcomes are derived through theory-based formula in one case and statistically for the remaining 12 SDGs (see Table 1). This was done through multi-country, cross-sectional multiple regression analysis and structural equation modeling (SEM). To capture direct and indirect associations, SEM was used for indicators that included mediators in their conceptual model.

Table 1. SDG Indicator Estimation Method

<table>
<thead>
<tr>
<th>Multiple Regression Analysis</th>
<th>Structural Equation Modeling</th>
<th>Theory-Based Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.1 Proportion of the population below the international poverty line</td>
<td>2.2.1 Prevalence of stunting among children under five years of age</td>
<td>3.7.2 Adolescent birth rate</td>
</tr>
<tr>
<td>2.1.2 Prevalence of moderate or severe food insecurity</td>
<td>3.1.1 Maternal mortality ratio</td>
<td></td>
</tr>
<tr>
<td>4.1.1 Proportion of children at the end of primary school achieving at least a minimum proficiency in reading</td>
<td>3.2.1 Under-five mortality rate</td>
<td></td>
</tr>
<tr>
<td>6.1.1 Proportion of the population using safely managed drinking water services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2.1 Proportion of the population using safely managed sanitation services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.1.1 Annual growth rate of real GDP per capita</td>
<td>11.1.1 Proportion of the urban population living in slums, informal settlements, or inadequate housing</td>
<td></td>
</tr>
<tr>
<td>8.2.1 Annual growth rate of real GDP per employed person</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.7.1 Proportion of children engaged in child labor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.1.1 Proportion of the urban population living in slums, informal settlements, or inadequate housing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Response variables for statistical testing were identified by applying two initial selection criteria to the 232 SDG indicators: (1) data availability and (2) the existence of a possible predictive or causal link between each response variable and either family planning use directly, or demographic variables more broadly (e.g., TFR, population size, or child dependency ratio). Twenty-one SDG indicators satisfied this initial criteria and were downloaded from the SDGs Indicator Global Database, produced by the UN Statistics Division.

Prior to statistical testing, a literature review (1) validated the assumed predictive or causal link to family planning or demography, and (2) identified the proximate and distal variables associated with each SDG indicator to inform the construction of both multiple regression and

---

2 SEM is a statistical technique used to evaluate the causal mechanisms through which independent variables affect a dependent variable. SEM uses a conceptual model and path diagram to capture relationships that include mediators, allowing variables to be both dependent and independent at different stages of the equation.
SEM models. After the literature review, 8 of the 21 SDG indicators were eliminated as possible outcome variables, and 13 response variables were selected for testing.

Guided by the literature review, data were collected on all relevant and available predictors, covariates, and mediators from 2000 on. Using a criterion of 70 observations (countries), HP+ collected data from compiler websites—for example, World Bank Open Data and UN Agency websites (e.g., UNICEF and UN Statistics Division)—as well as Demographic and Health Surveys, when possible. Depending on data availability for each indicator, we adjusted variables in the following ways:

1. Tested relationships using the same year for all predictor and response variables.

2. Tested relationships after calculating country averages for each indicator for 2000–2005, 2006–2010, and 2011–2015 to maximize the number of observations. To validate the consistency of relationships, predictors and response variables were tested for each time period. To derive the final statistical relationship, the latest period (2011–2015) was prioritized.

3. Tested relationships by using the latest data available for the indicator in cases of a large number of missing values and significant variation in indicator values in a short period of time (e.g., the prevalence of child stunting, reading proficiency at the end of primary school, and child labor). The predictor and response variable timeframes used were harmonized to reflect the same time period.

Initially, SEM was used to test 5 of the 13 SDG indicators: (1) the prevalence of stunting among children under five years of age, (2) MMR, (3) U5MR, (4) proportion of children at the end of primary school achieving minimum proficiency in reading, and (5) proportion of children engaged in child labor. The last two response variables were eliminated from SEM analysis because of the limited number of observations. Nine of the SDG indicators were estimated using multiple regression analysis; one indicator was computed based on a theory-derived formula (see Table 1).

For stunting, MMR, and U5MR indicators, SEM was estimated using the maximum likelihood method with a procedure that includes countries with missing values. This option allows all countries to be included in the model and all available information to be used for estimation, despite a degree of missing values. The maximum likelihood method assumes multivariate normality and requires a sample size close to 200 for optimal results. An extensive literature review of the causal mechanisms for each indicator—both proximate and distal—informed the construction tested paths. Decisions on whether to retain or exclude variables from each model were based on the significance of associations and the model fit. The latter was assessed using three tests: chi-square, root mean squared error of approximation (RMSEA), and comparative fit index (CFI), where a non-significant chi-square, RMSEA < 0.08, and CFI > 0.95 were desirable.

We used ordinary least squares estimation to test the predictive power of individual and combined sets of predictors on nine dependent variables. Because of heavily skewed distributions, some indicators (e.g., GDP per capita) were natural-log, cube, or square transformed. For each response variable, the model creation process proceeded as follows:

---

3 A predictor, also called an independent variable, is a variable that is manipulated in order to test the effect on a response, or dependent, variable. A covariate is a variable that is predictive of the response, or dependent, variable under study. A mediator is a variable that helps explain the relationship between a predictor and response variable. Therefore, a mediator variable is in the causal sequence from the predictor to the response.
1. Correlation matrices to examine the direction and strength of the relationship between each predictor and the response variable, as well as predictors to each other.

2. Bivariate testing, regressing the response variable on each predictor, individually.

3. Forward selection procedure beginning with the variable with the greatest theoretical relevance and best bivariate results. In some cases, the variable with the greatest theoretical relevance did not perform well in bivariate or multiple regression testing because of the data quality. Predictors were added one-at-a-time, based on the largest absolute t-value/t-statistic, while maintaining a p-value at the 5 percent level of significance and increasing $R^2$.

4. Regression diagnostics included checking for outliers, leverage, and influence; normality of residuals; and homoscedasticity.

Following statistical testing, the FP-SDGs Model was developed in Microsoft Excel to enable independent in-country use. Possible future values of some indicators have been capped so that they do not exceed a logical maximum. For a detailed description of the Excel model setup and explanation for running the model, see the User’s Manual section.

**Model Limitations**

This model was carefully constructed based on an extensive literature review, expert input, and application of rigorous statistical methods. Like all models, however, the FP-SDGs tool is based on a series of assumptions and it errs on the side of simplicity—allowing use by diverse audiences—rather than complexity. As a result, the model has limitations.

First, the statistical relationships in this model were estimated using international cross-sectional data assumed to (1) exist over time and (2) be applicable to any country in the model. These assumptions are central for model predictions and are used by previous family planning models, although they are unlikely to reflect reality for the distant future, or for each geography.

Second, the literature review shows that outcomes are multifaceted, having a variety of causes, and relationships occur at many levels (e.g., individual, household, community, and country). The use of the country as the unit of analysis has analytical limitations, preventing the use of variables that are predictive at other levels.

Data limitations were the greatest challenge for executing the estimations. In many cases, no data, or insufficient data, were available for the identified predictors. Thus, many possible predictors were excluded from the regression analysis or SEM because of a lack of data.

Finally, every model is a simplistic version of reality and does not include all the relevant factors. This should be considered when interpreting results, because model results should be considered estimates.

---

4 For example, the proportion of the population using safely managed drinking water and the proportion of the population using safely managed sanitation services were capped at 98 percent. In contrast, the prevalence of food insecurity could not fall below 0 percent.
User’s Manual

Overview

The FP-SDGs Model uses a blended software approach that combines an Excel model with the DemProj component of the Spectrum system of models. To complete a projection, five phases are needed:

- **Phase 1: Set up the basic model structure.** Following this step, the model is automatically loaded with default data for the model base year.

- **Phase 2: View and edit base year data inputs.** It is important for the user to provide missing data for any inputs that are not available in the default database. The user also has the option to replace default data using preferred international or local data sources.

- **Phase 3: Set policy goals for the model end year** for contraceptive use, as well as the optional policy variables (related to the sectors of education, governance, economic growth, agriculture, and health). These user-defined end-year values inform the SDG indicator projections.

- **Phase 4: Create projections** by linking the Excel model to the DemProj Spectrum module, which users must download.

- **Phase 5: View and validate results** by opening the Excel file after executing the population projections.

The Excel-based model has several additional pre-programmed worksheets for different functional areas. These additional worksheets are not intended for use.

**Phase 1: Set up the basic model structure**

To use the FP-SDGs Model, the user first needs to download and open the Excel file and enable the macros. To enable macros, after opening the Excel workbook (see Figure 6), select “Enable Content” in the yellow message bar with a shield icon.

**Figure 6. Enabling Macros in Excel**

Next, the user must use the drop-down menus in the **CONFIGURATION** worksheet to select the country of analysis, the timeframe (by indicating the base and end year), the base

---

year population source, and the type of model. Additionally, users have the option of setting minimum and maximum values—or caps—for select indicators (see Figure 7).

**Figure 7. CONFIGURATION Worksheet**

The user selects the country of analysis by navigating to the country list in Step 1. Next, the user chooses the base year and end year for the model. Base year options are currently 2015–2020, and end year options are 2030 (aligned with the SDGs timeframe) and 2050 (to allow additional time for improvements in SDG indicators to take effect). The FP-SDGs Model interpolates logistically (S-curve) between the base year and final year. The model produces results for every individual year over the projection period.

In Step 3, the user selects the source of their base year population data. If a user selects the Default population data option, the model will use the default data within DemProj (based on World Population Prospects data) to create future population projections. If the user selects Manual base year population entry, the BASE YEAR DATA INPUTS sheet will feature an empty table in columns G and H; the user then specifies the base year population by age and sex.
In Step 4, the user selects the type of model, choosing from two options: a model in which only family planning use changes over time; or one in which contraceptive use and socioeconomic components—related to education, governance, economic growth, agriculture, and other policy variables—change over time. This selection affects the information that will be displayed in the SET POLICY GOALS sheet.

In Steps 5 and 6, the user may either retain or edit existing caps—minimum and maximum values—for specific indicators (see Figure 7). For instance, a user may choose to assume that a country is unlikely to reach a maximum value, or 100 percent, for a specific variable given historical precedent or existing development conditions. In this example, the existing cap of 98 percent could be reduced to a lower value. Conversely, the cap could be raised to 100 percent.

After these selections are made on the CONFIGURATION sheet, the user must press the "Next" button to load the chosen parameters and proceed to the next page. A progress indicator box will appear during the loading process. Once all default data is loaded, the user will be taken to the subsequent page of the model.

**Phase 2: View and edit base year data inputs**

The BASE YEAR DATA INPUTS page includes 49 indicators across five thematic areas: contraceptive use, demography, health, education, and economics (see Figure 8). The user must view this page to both validate the auto-populated initial values as well as fill in any data gaps. Default source information for each variable is in the Data Source column in the worksheet, while exact indicator definitions are in the Definition column. If base year data are not available for the chosen country, the cells for variables without available data will appear in red with the value of 0.0. For the model to function, the user must enter missing base year values on the basis of alternate data sources or user assumptions. To replace existing data, alternate values may also be drawn from national sources or other references identified by the user and/or technical working group members, if applicable. Whenever data are edited or entered, it is recommended that the user edit the corresponding Data Source value to reflect the change. Cells edited/ altered by the user appear in light blue.

If a user makes a mistake and wants to override all the data entered on this page and revert to the default data, they can click the Reload All Default Data button. To clear all data, the user can click the Clear All Default Data button.

If the user selected Manual base year population entry in the CONFIGURATION sheet, the BASE YEAR DATA INPUTS sheet will feature an empty table in columns G and H; the user should specify the base year population by age and sex. The user should also enter the source data below the table.

After these revisions are made, the user must press the "Next" button to proceed to the next worksheet.
Phase 3: Set policy goals for the model end year

The next step in using the model is to establish policy scenarios (indicator goals) for the model end year in the **SET POLICY GOALS** sheet of the model. This enables the user to compare results based on different future scenarios.

In the **SET POLICY GOALS** sheet, the user must name each scenario; these names will be used in the **RESULTS** sheet to compare scenarios (see Figure 9).
The values that the user inputs for the model end year will depend on which of the following types of scenarios the user included in the final step in the **CONFIGURATION** sheet: (1) **Family planning scenarios only**, or (2) **Family planning and other policy variable scenarios**.

If the user selected **Family planning scenarios only** in the **CONFIGURATION** sheet:

- To compare the benefits of different family planning futures, and how they are expected to affect the SDG indicators included in the model, the user sets the desired family planning levels for up to three scenarios for the model end year.
- To do this, in the **SET POLICY GOALS** sheet, the user enters the following family planning information for married or in-union women:
  - Total CPR
    - Modern CPR
    - Traditional CPR
  - Unmet Need
- Scenario 1 is automatically populated with values from the **BASE YEAR DATA INPUTS** sheet. These values can be changed by the user at any time. They are auto-populated to provide an easy reference for setting scenarios, particularly if the user chooses to create a “constant” scenario (no changes from the base year).
- Definitions for these terms are in the **BASE YEAR DATA INPUTS** sheet, where the baseline values and data sources are also found.
- Demographic and Health Surveys, Multiple Indicator Cluster Surveys, Performance Monitoring and Accountability 2020 (PMA2020) surveys, and Track20 projections are good sources of data for this family planning information. Users may want to refer to past data and/or other countries’ data to determine realistic and ambitious scenarios for the model end year.
- Based on the information entered, the value for **Demand for Family Planning Satisfied with Modern Methods** will auto-calculate for each scenario. The calculation is modern contraceptive prevalence rate (mCPR) divided by total demand (total demand = total CPR + unmet need for contraception). Because the model’s default CPR and unmet need information is for married women, the auto-calculated demand for family planning satisfied with modern methods will also be for married women. The user may want to enter information for all women, in which case the auto-calculated indicator would refer to a broader sub-population of women.
The values for the optional policy variables noted in the section below are kept constant at their baseline values for all scenarios.

Under the **Select Model Type**, if the user selected **FP and other policy variable scenarios** in the **CONFIGURATION** sheet:

- In addition to the family planning information explained above, the user will also see end-year information for 12 **OPTIONAL POLICY VARIABLES** for each scenario.
- These variables are all defined in the **BASE YEAR DATA INPUTS** sheet, where the baseline values and data sources can also be found (see also Annex A). Additional information about the Global Competitiveness Index (GCI) variables is also in the Gross Domestic Product section of the Methodology (within Module 3 on Economics).
- For ease of developing scenarios, the “Scenario 1” column of the **SET POLICY GOALS** sheet is populated with either default or user-replaced data from the **BASE YEAR DATA INPUTS** sheet (see Figure 10).
- Users may want to refer to past data and/or other countries’ data to determine realistic and ambitious scenarios for the model end year for each indicator.

**Figure 10. SET POLICY GOALS Sheet with Family Planning and other Policy Variable Scenarios**

To view the differential and combined impacts of family planning and other policy investments, the user may consider creating the following scenarios:

- **Scenario 1**: “constant” scenario, in which no changes are made to any variables from the base year of the model. The results inform the user of what can be expected for SDG

---

Another option is to replace the “constant” scenario with a “business-as-usual” scenario, in which recent changes, over time, are continued into the future at the same rate of change as in the past.
indicators if the status quo continues (i.e., with no changes in the current policy environment).

- **Scenario 2**: “socioeconomic and governance improvements” scenario, in which aspirational end year changes are applied to each of the 12 multisectoral policy variables listed above. Family planning metrics may be held constant from scenario one.

- **Scenario 3**: “combined family planning, socioeconomic, and governance improvements” scenario, in which the user replicates the changes of scenario 2, while increasing achievement/progress in the family planning metrics. This scenario typically produces the most pronounced improvements in results.

- Alternatively, scenario 2 can include conservative improvements for governance and socioeconomic, and family planning variables; while scenario 3 includes more aspirational improvements for these variables.

- Another option is to include aspirational governance and socioeconomic goals for all three scenarios, while varying the family planning levels across scenarios, with scenario 1 showing little to no improvements in family planning indicators, scenario 2 showing conservative family planning improvements, and scenario 3 showing more aspirational family planning improvements.

The end value for these variables is often set at an optimistic yet achievable level, balancing the significant changes that are likely needed with the feasibility of such changes given the available and required resources. In some cases, national development plans, sector policies, and other documents will offer specific goals and targets that can be adopted for the policy scenarios for this model. Figure 11 shows illustrative policy scenarios for Malawi.

**Figure 11. Illustrative Policy Scenarios for Malawi**
Phase 4: Create projections

The next step is to run the RAPID Transfer tool for the SDG model in Spectrum in order to perform population projections. This step pulls outputs from the FP-SDGs Model for TFR and male and female life expectancy and shares them with the DemProj module of Spectrum, where they are used as inputs for population projections. Several key demographic variables are then exported back to the Excel FP-SDGs Model from DemProj. The two models automatically complete almost all the work for this step; the user only needs to set up the communication between them.

First, to open Spectrum, the user should search for Spectrum on the user’s computer and click on the program icon (see Figure 12). If not already installed, Spectrum can be downloaded free of charge, with the accompanying Country Data Pack. They are available on the Avenir Health website, at www.avenirhealth.org/software-pc.php.

When the Spectrum dialogue box is displayed, the user clicks on the option to Run Spectrum (see Applications and Tools) (see Figure 13). Before proceeding, the user should save and close their FP-SDGs Model file in Excel. Also, when using the RAPID Transfer tool in Spectrum, the user cannot open Spectrum files. In Spectrum, the user should do the following (see Figure 14):

- Click on the Tools tab.
- Click on the More tools icon.
- Scroll down to the External group and click on the RAPID Transfer icon.
- Select the SDG model option.
- Select the country.
- Click Select SDG File and navigate to where the user’s FP-SDGs Excel file is saved.
- Click the Process button.
- If a window appears asking if you want to download default demographic data, click Yes and then OK.
- When a window appears asking if you want to save the workbook, click Yes.
  - If an error message says that the file must be closed while it is being processed, this is because the Excel FP-SDGs Model file is open. Close the Excel file for the interaction with Spectrum to work.
- When the status bar shows 100 percent, click Close.

If the “Manual base year population entry” option was selected on the CONFIGURATION page, then the age- and sex-disaggregated population data that the user entered will also be shared with DemProj.
To close Spectrum, click the X at the top right of the two windows.

Click Yes when the warning appears: “Data have changed, save all projections?” (Three projection files will be generated—one for each scenario. They will be saved automatically in the same location as the FP-SDGs Model Excel file. Do not do anything else with these new files.)

**Figure 13. Running Spectrum**

**Figure 14. Initial Steps in Spectrum**
Phase 5: View and validate results

After completing the previous step, the user should open their FP-SDGs Model file in Excel. In the RESULTS sheet, from the drop-down menu, the user can select the SDG indicator model result they want graphed (see Figure 15). The graph will compare the scenarios set in the SET POLICY GOALS sheet.

Figure 15. Selecting SDG Indicator to Graph

Figure 16 shows illustrative results graphs from Malawi, based on the calibration displayed in Figure 11.

Figure 16. Illustrative Results Graphs from FP-SDGs Model (Malawi)

Also, in the RESULTS tab, the user can see summary results (in a table located below the graph) for all indicators at one time (see Figure 17). This table contains the base year indicator values and the end year indicator values for all scenarios in the columns labeled Scenario Results Summary. In the table, on the right, the user will see a pair of columns labeled Scenario Comparison: Absolute Change. These columns compare scenario 2 to scenario 1, and scenario 3 to scenario 1. To the right of these columns are two columns labeled Scenario
**Comparison: Percentage Change.** These columns also compare scenario 2 to scenario 1, and scenario 3 to scenario 1, but they show the percentage change in values (instead of the absolute change).

Figure 17. Illustrative Summary Results Table (Malawi)

<table>
<thead>
<tr>
<th>STEP 11: VIEW DATA BY SDG AREA</th>
<th>SCENARIO RESULTS SUMMARY</th>
<th>SCENARIO COMPARISON: ABSOLUTE CHANGE</th>
<th>SCENARIO COMPARISON: PERCENTAGE CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2060</td>
<td>Constant CPRI</td>
<td>Incremental CPRI</td>
<td>Ambitious CPRI</td>
</tr>
<tr>
<td>Goal 1. Good Health and Well-Being</td>
<td>39</td>
<td>26.7</td>
<td>26.5</td>
</tr>
<tr>
<td>Goal 2. No Poverty</td>
<td>41</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Goal 3. Quality Education</td>
<td>49</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Goal 4. Clean Water and Sanitation</td>
<td>61.3</td>
<td>65.6</td>
<td>67.3</td>
</tr>
<tr>
<td>Goal 5. Decent Work and Economic Growth</td>
<td>1.8</td>
<td>1.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Goal 6. Reduced Inequalities</td>
<td>1.4</td>
<td>1.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Goal 7. Opportunity to Learn</td>
<td>51.6</td>
<td>42.3</td>
<td>41.0</td>
</tr>
<tr>
<td>Goal 8. Contraception Services</td>
<td>37.4</td>
<td>13.4</td>
<td>13.0</td>
</tr>
</tbody>
</table>
Module 1: Demographics

Overview
The demographic module within the FP-SDGs Model is primarily used to create population projections through DemProj. Outcomes, subsequently, feed into the estimations of the health, economics, and development modules. Of the five estimations in this module, four are from/identical to those of DemDiv. For a full account of each relationship, please refer to the Technical Guide to the DemDiv Model.

Estimations

Total fertility rate
The TFR is the heart of the demographic model; it calculates the number of annual births, which informs population growth and age structure. Following the structure of the DemDiv model, the FP-SDGs Model adapts the Bongaarts (1978) proximate determinants model to project TFR. Three variables directly affect the TFR, which the user enters, and it is indirectly affected by girls’ education, by marriage. Under Bongaarts’ framework, fertility is computed as \( TFR_t = Cm_t \times Cc_t \times Ci_t \times Ca_t \times TF \) (see below):

- The index of marriage (Cm) is the percentage of women of reproductive age who are married or in-union. Future values are derived using the estimated relationship between education and marriage (see Table 2).
- The index for contraception (Cc) is calculated as a function of the CPR and method mix. CPR is interpolated, over time, using an “s-curve” pattern—or logistic model—to reach user-defined scenario values. The index is calculated from the prevalence (prev) and effectiveness of modern (m) and traditional (trad.) methods as: \( Cc = 1 - 1.08 (prev_m \times \text{effectiveness}_m) + (prev_{trad.} \times \text{effectiveness}_{trad.}) \). We assume 95 percent and 50 percent effectiveness for mCPR and traditional methods, respectively.
- The index of postpartum insusceptibility (Ci) reflects the duration of a woman’s temporary inability to conceive because of postpartum amenorrhea from breastfeeding and postpartum abstinence. The insusceptibility index (Ci) is calculated as: \( Ci = 20.0 / (18.5 + \text{period of postpartum insusceptibility in months} \). In the model, it is assumed to stay constant over time.
- The FP-SDGs Model omits the index of induced abortion (Ca) and calculations were adjusted (solved) accordingly.
- Total fecundity (TF) represents the biological maximum number of children the average woman might have in her lifetime. TF is not a model input; instead, the model solves for TF based on its calculations of the other proximate determinants.

In summary, the model calculates TFR by retaining Cc, Ci, Cm, and TF. Cc and Ci are exogenous and user controlled. TF is solved for using the TFR equation. Cm is endogenous and calculated by the model, as described below, as a function of girls’ education.

---

8 Bongaarts’ original framework includes four indices and total fecundity; the index of pathological sterility was added at a later date and is excluded from our calculations.
**Percentage of women married or in-union**

The percentage of women who are married or in union, described above as Cm, is an important proximate determinant of fertility because it reflects the share of women who are, presumably, regularly sexually active (Bongaarts, 1978). In the FP-SDGs Model, as in DemDiv, this variable is modeled as a function of female education. Increased educational attainment can lead to delays in the age of marriage as women and girls stay in school longer (Aryal, 2007; Islam and Ahmed, 1998). In this way, female education indirectly impacts fertility and, therefore, population dynamics. For a full account of the estimation process, please refer to the *Technical Guide to the DemDiv Model*. We re-estimated the original DemDiv equation using the latest data.

**Table 2. Estimated Equation for Percentage of Women Married or In-Union**

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of women married or in-union</td>
<td>Mean years of female education</td>
<td>-1.07</td>
<td>-5.89</td>
<td>.20</td>
<td>145</td>
</tr>
</tbody>
</table>

**Female education**

In addition to the link between female education, marriage, and fertility, as described above, the model includes a feedback loop to capture the effect of fertility on female education (see Table 3). This concept is again from the DemDiv model, but is re-estimated using recent data.

**Table 3. Estimated Equation for Female Education**

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean years of female education</td>
<td>TFR</td>
<td>-.34</td>
<td>-15.1</td>
<td>.61</td>
<td>145</td>
</tr>
</tbody>
</table>

**Female life expectancy**

Female life expectancy is a high-level health indicator that synthesizes mortality at all ages at one point in time. In the FP-SDGs Model, as in DemDiv, this variable is modeled as a function of the under-5 mortality rate (see Table 4). For a full account of the estimation process, please refer to the *Technical Guide to the DemDiv Model*.

**Table 4. Estimated Equation for Female Education**

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(female life expectancy)</td>
<td>Ln(U5MR) when U5MR &lt; 50.9</td>
<td>-.59</td>
<td>-18.3</td>
<td>.950</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>Ln(U5MR) when U5MR &gt; 50.9</td>
<td>-.287</td>
<td>-22.5</td>
<td>.950</td>
<td>196</td>
</tr>
</tbody>
</table>

**SDG indicator 3.7.2: Adolescent birth rate**

The adolescent birth rate is the annual number of births per 1,000 females ages 15–19 (UN Statistics Division, 2017), and is also called the age-specific fertility rate (ASFR) for the 15–19 age group. In the model, the adolescent birth rate is derived from the TFR and is an output from DemProj. Specifically: Adolescent birth rate at time \( t \) = \( TFR_t * ASFR\_{15–19}\_{year-olds} / 100 \).
Module 2: Health

Overview

The health module within the FP-SDGs Model computes four key SDG indicator outcomes. Multiple regression is used to estimate food insecurity. The prevalence of stunting, the MMR, and the U5MR are estimated using structural equation modeling to address mediators in the relationships. All estimations use country-level variables, therefore, other causes at the individual level (e.g., birth order and birth weight) are not considered in the associations.

Estimations

SDG indicator 2.1.2: Prevalence of moderate or severe food insecurity

The prevalence of moderate or severe food insecurity is an indicator from the Food Insecurity Experience Scale of the Food and Agriculture Organization of the United Nations.\(^9\) The scale is computed from an eight-question dichotomous (yes/no) survey gauging access to adequate quantities of nutritious food during a period of one year (UN Statistics Division, 2017).

Food security depends on relative stability, over time, for three factors: (1) food availability, (2) economic and physical access to food, and (3) food utilization/uptake (FAO, 2008). Food availability is linked to having an adequate food supply, including through yield-enhancing technology, sustainable agriculture policy, and functioning markets (FAO et al., 2015; World Bank, 2015). Despite increases in food production over time, higher fertility increases food demands, and it may constrain supply factors under certain conditions (Bongaarts, 1996). Moreover, fluctuations in household-, community-, and national-level conditions—periods of conflict, weather-related shocks, rising food prices, unemployment, etc.—may periodically lead to food insecurity (FAO, 2008; FAO et al., 2015).

Food access depends on individuals' financial assets and the policy environment, such as income and the existence of pro-poor programs like food distribution schemes, cash transfers, and school-feeding projects (FAO et al., 2015; World Bank, 2015). Finally, utilization—commonly understood as the way the body processes nutrients—is the result of feeding practices, food preparation, diversity of the diet, and other factors within the household (FAO, 2008).

Because of data limitations, only 12 relevant predictor indicators were available for testing using multiple regression. Tested variables included available arable land; a measure of environmental sustainability; crop, food, and cereal production; a measure of social protection; and minimum dietary diversity; among others. Individually and combined, cereal production per capita, income per person, and TFR had a statistically significant impact on food insecurity, accounting for 71 percent of the variation in the response variable (see Table 5).

\(^9\) Food security is a state “when all people, at all times, have physical, social, and economic access to sufficient, safe, and nutritious food that meets their dietary needs and food preference for an active and healthy life” (FAO, 2008).
Table 5. Food Insecurity Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence of moderate or severe food insecurity</td>
<td>Ln(cereal production per capita)</td>
<td>-2.32</td>
<td>-3.02</td>
<td>.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GDP per capita)</td>
<td>-5.33</td>
<td>-4.95</td>
<td>.000</td>
<td>.71</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Total fertility rate</td>
<td>7.70</td>
<td>7.04</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SDG indicator 2.2.1: Prevalence of stunting among children under five years of age**

Stunting among children under-five—or impaired growth by age—is defined as height-for-age that is more than two standard deviations below the median of the World Health Organization (WHO) Child Growth Standards (UN Statistics Division, 2017). Inadequate feeding practices and childhood disease are the two most immediate causes of stunting. After birth, early initiation of breastfeeding and exclusive breastfeeding bolsters infant immunity and prevents gastrointestinal infections, including those that could lead to nutrient depletion and stunting. After six months, breastmilk should be complemented by nutrient-rich foods—including increasing the amount and variety, over time, to meet minimum dietary diversity standards\(^\text{10}\)—to ensure normal growth (WHO, 2014). Moreover, preventing and treating severe illnesses—like diarrhea and malaria—is essential for promoting linear growth (WHO, 2014). Key interventions include prompt healthcare-seeking behavior and treatment of disease (with oral rehydration and continued feeding, antimalarial drugs, etc.), reducing environmental contaminants, and improving hygienic practices (WHO, 2014; Devlin, 2012).

Poor maternal nutrition before, during, and after pregnancy impairs fetal development and also contributes to stunting, requiring interventions like folic acid supplementation, balanced energy-protein supplementation, and multiple micronutrient supplementation (IFPRI, 2016). Stunting is also impacted by the timing and number of pregnancies/births. Specifically, the risk of child stunting decreases with increased time between the preceding birth and the conception of the next child; the optimal spacing period is at least 30 months. Stunting outcomes are also most common when a mother is under the age of 18 because of competition for nutrients between the still-growing mother and her fetus, as well as with increased parity (four or more children) (Rutstein and Winter, 2014; WHO, 2014). Many more distal drivers also affect linear growth, including poverty, because of its impact on food accessibility (adequate feeding practices) and maternal education, which affects healthcare-seeking behavior and the correct treatment of childhood illness (WHO, 2014).

SEM was executed for stunting based on these relationships and the available data. We began by testing six separate path models, one for each of the possible mediators. Each model tested one of the key proximate drivers of stunting—three measures of feeding practices and three measures of childhood disease\(^\text{11}\)—along with exogenous variables like log GDP per capita, female

---

10 Refers to feeding the child food from at least four food groups originating from seven categories: grains, roots, and tubers; legumes and nuts; dairy products (milk, yogurt, cheese); flesh foods (meat, fish, poultry, liver/organ meat); eggs; vitamin A-rich fruits and vegetables; and other fruits and vegetables.  
11 Minimum acceptable diet, minimum dietary diversity, exclusive breastfeeding, malaria treatment, diarrhea treatment, and prevalence of anemia.
mean years of education, food production, crop production, TFR, adolescent birth rate, and others. Exogenous variables were treated as direct predictors of the SDG indicator, and/or as covariates of mediators. The paths for disease treatment were excluded from the composite SEM model because no associations were found in the preliminary, separate analysis. Minimum dietary diversity was retained as a mediator, and the model explained 85 percent of the variation in stunting. The decision to exclude additional paths, exogenous variables, and covariates was made based on the significance or fit of the model. Table 6 presents the path analyses for the final stunting SEM model. For the related path diagram, see Annex B.

<table>
<thead>
<tr>
<th>Response Variables</th>
<th>Covariates</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean years of female education</td>
<td>1.73</td>
<td>0.055</td>
<td>.78</td>
</tr>
<tr>
<td></td>
<td>TFR</td>
<td>-7.48</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GDP per capita)</td>
<td>6.57</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Stunting</td>
<td>Minimum dietary diversity</td>
<td>-0.17</td>
<td>0.002</td>
<td>.76</td>
</tr>
<tr>
<td></td>
<td>Ln(GDP per capita)</td>
<td>-3.07</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean years of female education</td>
<td>-1.37</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
</tbody>
</table>

Overall model: — — — .85

Goodness of fit: Chi2(1): 0.97; RMSEA: 0.00; CFI: 1.00.
Co-variances between exogenous variables are established as a standard procedure when utilizing the maximum likelihood for missing values method. They are not presented here because they do not alter the coefficients for predicting the outcome.

**SDG indicator 3.1.1: Maternal mortality ratio**

The MMR is defined as the number of maternal deaths per 100,000 live births during the same time period. A maternal death refers to mortality attributed to any cause related to or aggravated by pregnancy or its management during a specific period (UN Statistics Division, 2017). Direct causes include obstetric factors like hemorrhage, while indirect causes of death result from pre-existing conditions—or those that develop during pregnancy—and are aggravated by its state, like malaria or anemia (WHO et al., 2015). Driving these direct and indirect causes are contraceptive use, pregnancy care, maternity care, as well as more distal socioeconomic components.

Pregnancies that are more likely to result in life-threatening complications include those that occur at too young or too old maternal ages, are spaced too closely, are at high parities, or would have ended in an unsafe abortion. These pregnancies are classified as being of greater-than-average risk to survival (high-risk fertility behavior). Contraceptive use not only directly reduces the risk of maternal death by decreasing exposure to pregnancy, but it also reduces the average risk per pregnancy by helping distribute births from the high- to low-risk categories (Cleland et al., 2012; HPI and USAID, 2008; Ahmed et al., 2012).

Availability of high-quality pregnancy and maternity care is essential for maternal survival. Antenatal care is important for detecting, treating, and preventing chronic conditions like anemia, which might cause complications (Carroli et al., 2001). Skilled assistance during childbirth is essential for managing childbirth safely, while access to basic and comprehensive emergency obstetric care services is critical for managing the complications that account for
most maternal deaths. Finally, skilled care during the 24–48 hours following delivery is an essential strategy for detecting and treating postnatal complications (Campbell et al., 2006).

Distal socioeconomic factors impact a patient’s ability to seek and access health services from pre-pregnancy to childbirth. These factors—many of which are embedded in the “three delays model” (Thaddeus and Maine, 1994)—include poor maternal education (Karlsen et al., 2011); poverty or financial barriers to receiving care (Thaddeus and Maine, 1994); and the laws, policies, program budgets, and social protection schemes that affect the availability, accessibility, acceptability, and quality of services (WHO, n.d.).

SEM was executed for the MMR based on these relationships and available data. We began by testing five separate path models. Each model tested one of the key proximate drivers of maternal death as a possible mediator—skilled birth attendance, postnatal care, high-risk fertility behavior, anemia during pregnancy, and short birth spacing—along with exogenous variables selected based on a literature review (e.g., mean years of female education, gender equality, government health expenditure, gender equality, antenatal care, and others). Exogenous variables were treated as explanatory variables of the mediator and/or as direct predictors of the outcome.

The high-risk fertility behavior and short birth spacing paths were excluded from the composite model because of extensive missing values. As a result, TFR and the adolescent birth rate (both exogenous) were included in the model as proxies for birth spacing and the risks associated with early childbearing, respectively. The path for postnatal care was excluded from the composite SEM model because of its high correlation with antenatal care and skilled birth attendance, and non-association with the MMR when these variables were entered as controls. Skilled birth attendance was retained as the only mediator, affected by antenatal care, female mean years of education, and income per person. The model explained 83 percent of the variation in the MMR. The decision to exclude additional paths, exogenous variables, and covariates was made according to the significance or the fit of the model. Table 7 presents the path analyses for the final MMR SEM model. For the related path diagram, see Annex B.

### Table 7. Path Analyses for MMR

<table>
<thead>
<tr>
<th>Response Variables</th>
<th>Covariates</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skilled birth attendance</td>
<td>Antenatal care</td>
<td>.28</td>
<td>&lt; 0.001</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>Mean years of female education</td>
<td>2.44</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GDP per capita)</td>
<td>2.88</td>
<td>&lt; 0.008</td>
<td></td>
</tr>
<tr>
<td>Maternal mortality ratio</td>
<td>Skilled birth attendance</td>
<td>-3.48</td>
<td>&lt; 0.001</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>TFR</td>
<td>70.31</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adolescent birth rate</td>
<td>1.32</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Overall model</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.83</td>
</tr>
</tbody>
</table>

Goodness of fit: Chi²(6): 14.96; RMSEA: 0.09; CFI: .98. Co-variances between exogenous variables are established as a standard procedure when utilizing the maximum likelihood for missing values method. They are not presented here because they do not alter the coefficients for predicting the outcome.
**SDG indicator 3.2.1: Under-five mortality rate**

The U5MR is the probability of a child dying before reaching the age of five if subject to age-specific mortality rates of that period, expressed per 1,000 live births (UN Statistics Division, 2017). Approximately half of all under-five deaths occur during the neonatal (28 days postpartum) period, driven by pre-term complications and intrapartum events, like asphyxia and infections. Beyond the neonatal period through age five, the main causes of death include pneumonia, diarrhea, and other causes originating during the perinatal period (Liu et al., 2016). Driving these direct causes of death are factors related to availability, quality, and uptake of preventive and curative child health services, environmental conditions, and more distal socioeconomic circumstances.

The availability of quality healthcare services is essential for addressing the pre-term, intrapartum, and postpartum complications leading to the high burden of neonatal mortality. These health services include antenatal care for detecting and managing pre-existing conditions; skilled birth attendance; and comprehensive emergency obstetric care for treating complications like asphyxia, breastfeeding initiation, and infection prevention, as well as postnatal checkups (UN IGME, 2017). Beyond the neonatal period, combating under-five mortality requires the availability of vaccinations, proper nutrition, and prompt management of childhood illnesses—for example, diarrhea, malaria, and pneumonia—through oral rehydration therapy and treatment with antimalarial drugs and antibiotics (WHO, 2017a).

Like maternal deaths, under-five mortality is also impacted by fertility behavior, particularly the timing of pregnancies, although the exact mechanism through which this effect occurs continues to be studied (Conde-Agudelo et al., 2012). Specifically, very short or very long birth-to-conception or birth-to-birth intervals increase the risk of child mortality (Cleland et al., 2012; Rutstein and Winter, 2014). Children born to young mothers also have an elevated risk of dying in the first year of life (Rutstein and Winter, 2014).

Moreover, reducing household air pollutants, as well as improving access to safe water, sanitation, and hygiene are essential for creating an enabling environment for child health (Liu et al., 2016). Finally, distal socioeconomic factors impact a caretaker’s ability to seek and access pregnancy, birth, and child healthcare services; socioeconomic factors that can negatively affect under-five mortality include poor maternal education (Mosley and Chen, 1984) and poverty or financial barriers to receiving care (UNICEF, 2015).

SEM was executed for the U5MR based on these relationships and available data. We began by testing seven separate path models for each proximate, health service-related cause: the hypothesized mediators. These variables were skilled birth attendance, postnatal care, tetanus immunization, diarrhea treatment, malaria treatment, maternal anemia during pregnancy (proxy for maternal nutrition), and pneumonia healthcare-seeking behavior. Additional exogenous variables were tested in each path model based on a literature review; these included mean years of female education, TFR, log GDP per capita, and others. Exogenous variables were treated as explanatory variables of the mediators and/or as direct predictors of the outcome.

Of those variables tested, most pathways were excluded because of non-associations with the outcome or worsening model fit. Skilled birth attendance and tetanus immunization were retained as mediators; the overall model explained 89 percent of the variation in the U5MR. Table 8 presents the path analyses for the final U5MR SEM model. For the related path diagram, see Annex B.
Table 8. Path Analyses for Under-Five Mortality Rate

<table>
<thead>
<tr>
<th>Response Variables</th>
<th>Covariates</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skilled birth attendance</td>
<td>Antenatal care</td>
<td>.28</td>
<td>&lt; 0.001</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>Mean years of female education</td>
<td>2.53</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GDP per capita)</td>
<td>2.67</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Tetanus immunization</td>
<td>Intercept</td>
<td>75.88</td>
<td>—</td>
<td>.40</td>
</tr>
<tr>
<td></td>
<td>Adolescent birth rate</td>
<td>-0.05</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Antenatal care</td>
<td>.19</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Under-five mortality rate</td>
<td>Intercept</td>
<td>77.74</td>
<td>—</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>Skilled birth attendance</td>
<td>-0.35</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tetanus immunization</td>
<td>-0.42</td>
<td>.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFR</td>
<td>9.55</td>
<td>&lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adolescent birth rate</td>
<td>.15</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean years of female education</td>
<td>-1.54</td>
<td>.0031</td>
<td></td>
</tr>
<tr>
<td>Overall model</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.89</td>
</tr>
</tbody>
</table>

Goodness of fit: Chi²(8): 17.17; RMSEA: 0.07; CFI: .98. Co-variances between exogenous variables are established as a standard procedure when utilizing the maximum likelihood for missing values method. They are not presented here because they do not alter the coefficients for predicting the outcome.
Module 3: Economics

Overview

Except for estimating one SDG indicator, the entire economics module is from the DemDiv model. In DemDiv, this module integrates demographic outputs and several policy variables to project the GDP. For a full account of the GDP modeling process, please refer to the *Technical Guide to the DemDiv Model*. The following is a basic overview of the estimations required to project GDP. Breaking from DemDiv, we also compute the SDG indicator 1.1.1—the proportion of the population living below the international poverty line—using multiple regression.

Estimations

**SDG indicator 1.1.1: Proportion of the population below the international poverty line**

This indicator is defined as the percentage of the population living on less than $1.90 a day at 2011 international prices (UN Statistics Division, 2017). Many factors are associated with poverty defined in this traditionalist view, and these cut across multiple tiers (e.g., regional, community, household, and individual level) (Haughton and Khandker, 2009). For this model, we focus on three categories of drivers: pro-growth correlates, pro-poor correlates, and demographic-institutional factors.

Economists have observed that poverty tends to fall as economies—specifically average incomes—grow (Dollar and Kraay, 2002). Using data from developing countries, Kraay (2006) demonstrated that most of the variation in changes in poverty are attributed to growth in average incomes. Others suggest that the poor do not benefit equally from income growth (Ravallion, 2001) and that the level of inequality affects income elasticity (Fosu, 2010). Demographic factors, including TFR, have also shown to increase poverty through several mechanisms, including by skewing the distribution of income growth against the poor (Eastwood and Lipton, 2001). Economists tested the impact of various policy and institution variables on poverty—like trade openness, inflation, school enrollment, and government expenditures on health and education—although with mixed effects (Dollar et al., 2016; Rodrik, 2000).

Building on this evidence, we tested 27 predictors on the outcome (e.g., income per person, trade openness, labor force participation rate, unemployment), pro-poor measures (Gini coefficient, public spending on health and education, social protection scores, land reforms, etc.), and demography (TFR, population growth rate, life expectancy, etc.). The final model included GDP per capita, the Gini coefficient, and TFR, explaining 83 percent of the variation in the proportion of the population below the international poverty line (see Table 9).

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>$R^2$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sqrt(Proportion of the population below the international poverty line)</td>
<td>Ln(GDP per capita)</td>
<td>-.885</td>
<td>-5.25</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gini coefficient</td>
<td>.073</td>
<td>5.35</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFR</td>
<td>.794</td>
<td>5.90</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9. Population below the Poverty Line Estimation
**Gross Domestic Product**

GDP is used in the computation of two SDG outcome indicators: (1) the annual growth rate of real GDP per capita and (2) the annual growth rate of real GDP per employed person. Building on the Cobb-Douglas (1928) production function, three components are used in the projection of GDP: capital stock, human-capital augmented labor, and TFP. Capital stock is projected in the model by estimating investment per working-age adult. Human-capital augmented labor is derived through an estimation of employment combined with an education parameter set by the user. Both investment and employment are derived through multiple regression, integrating results from the demographic module, as well as economic policy variables. TFP is based on the regression outcome of three economic variables only.

The policy variables used in these computations are from *The Global Competitiveness Report*, published annually by the World Economic Forum (Schwab, 2012). The report measures the relative competitiveness of the world’s economies through its Global Competitive Index (GCI). This index is computed from a comprehensive database of more than 100 indicators arranged in 12 pillars of microeconomic and macroeconomic national competitiveness, defined as “the set of institutions, policies, and factors that determine the level of productivity of a country” (Schwab 2012, p. 4).

Three indicators are predictors of investment per working-age adult in the model, also referred to as capital formation per capita (see Table 10): real GDP per working-age adult, ratio of the population ages 15 and older to the total population, and a GCI variable on financial market efficiency. The GCI variable is an index of factors relating to access to financial services, loans, and venture capital. Regardless of other factors, individuals and businesses are unlikely to invest in an economy unless they can do so easily and without excessive costs.

### Table 10. Investment Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(investment/ working-age adult)</td>
<td>Ln(GDP per working-age adult)</td>
<td>.875</td>
<td>30.63</td>
<td>.97</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>Ln(working-age/total population)</td>
<td>.77</td>
<td>2.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GCI 8A: financial market efficiency)</td>
<td>.352</td>
<td>2.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P-values are not available for DemDiv statistical tests, but all were significant at the .05 level or less.

Employment is a function of the growth rate of the working age population, the growth rate of real GDP, and a GCI variable on labor market flexibility (see Table 11). The GCI subpillar is formed from indices of cooperation in labor-employer relations, flexibility of wage determination, hiring and firing practices, redundancy costs, and the extent and effect of taxation—each likely to affect employment growth. The results of the employment equation are then combined with mean years of education for both men and women to project the efficiency of the labor force, a direct input into the GDP production function.
Table 11. Employment Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(employment)</td>
<td>Ln(growth rate of working-age population)</td>
<td>.682</td>
<td>8.63</td>
<td>.64</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>Ln(growth rate of GDP)</td>
<td>.483</td>
<td>4.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GCI 7A: labor market flexibility)</td>
<td>.593</td>
<td>1.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P-values are not available for DemDiv statistical tests, but all were significant at the .05 level or less.

Finally, TFP measures how efficiently economic inputs are being used. In the model, TFP is a function of three GCI variables (see Table 12), selected based on the literature. First, GCI subpillar 1A on public institutions measures property rights (including intellectual property), division of powers, corruption, regulatory burdens, transparency, waste in government spending, and public safety. The second GCI variable is subpillar 9B on information and communication technology use, which includes indicators on Internet use, connectivity, bandwidth, and mobile phone subscriptions. The third GCI variable measures trade openness, an indicator measuring imports as a percentage of GDP under GCI pillar 6, Goods Market Efficiency.

Table 12. TFP Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(TFP)</td>
<td>Ln(GCI 1A: public institutions)</td>
<td>.623</td>
<td>3.10</td>
<td>.84</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Ln(GCI 9B: information and communication technology use)</td>
<td>1.187</td>
<td>14.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ln(GCI 6.14: imports as a percentage of GDP)</td>
<td>-.219</td>
<td>-3.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P-values are not available for DemDiv statistical tests, but all were significant at the .05 level or less.

**SDG indicator 8.1.1: Annual growth rate of real GDP per capita**

Based on the methodology described in the Gross Domestic Product section, the FP-SDGs Model projects a real GDP value for each user-defined year (e.g., 2020–2050). GDP is then divided by the total projected population per year—an output of the demographics module. Using these per capita GDP values, growth rates/percentage change is computed for each consecutive two-year period.

**SDG indicator 8.2.1: Annual growth rate of real GDP per employed person**

Based on the methodology described in the Gross Domestic Product section, the FP-SDGs Model projects a real GDP value for each user-defined year (e.g., 2020–2050). GDP is then divided by the projected employed population. Using these values, growth rates/percentage change is computed for each consecutive two-year period.
Module 4: Development

Overview

The development module within the FP-SDGs Model computes five indicator outcomes: the proportion of children at the end of primary school achieving at least a minimum proficiency level in reading; proportion of the population using safely managed drinking water services; proportion of the population using safely managed sanitation services; proportion of children engaged in child labor; and proportion of the urban population living in slums, informal settlements, or inadequate housing. Two of these indicators (SDG indicator 4.1.1 on reading proficiency and SDG indicator 8.7.1 on child labor) were selected for testing using SEM, but the limited number of observations prevented its use. All outcomes are estimated using multiple regression.

Estimations

**SDG indicator 4.1.1: Proportion of children at the end of primary school achieving at least a minimum proficiency level in reading**

SDG indicator 4.1.1 measures the “proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex” (United Nations, 2018, pg. 5). For the model, the indicator has been restricted to the end of primary school because the evidence is more robust for this grade level compared to others.

The measurement of learning outcomes by subject has not received much attention in the academic literature beyond in high-income countries where this data are more frequently collected. Additionally, countries continue to use differential benchmarks/standards and assessments, challenging comparability across countries. More commonly used education indicators include enrollment, grade progression, literacy, and completion of schooling.

The factors most closely linked to reading achievement occur at the level of the student, teacher, school, and family. A child’s working memory and attention control play a key role in developing literacy and mathematics skills and achievement (Welsh et al., 2010), as do general attitudes toward school and school life (Marks, 1998). Teacher behaviors—like classroom management and the diversity of teaching strategies—have one of the most robust relationships with children’s achievement in school, as do teachers with higher self-efficacy and those with better qualifications (Muijs and Reynolds, 2002; Fuchs and Woessmann, 2004).

Generally, schools with strong institutional practices—like exit exams and other forms of routine monitoring of student learning, effective administrative leadership, and quality instructional material—have higher student performance (Hanushek, 1992; Fuchs and Woessmann, 2004). Finally, strong family literacy practices, access to reading material in the home, parent educational achievement, and quality of parent-teacher relationships positively impact a child’s learning outcomes (Dearing et al., 2006; Senechal and LeFevre, 2002; Hughes and Kwok, 2007; Fuchs and Woessmann, 2004).

Fertility behavior and a student’s educational achievement are also linked. Family size affects parents’ human capital investments in their children—the larger the family size, the smaller the investment in child education, which is related to a lower quality of education and lower achievement in school (Becker, 1960; Hanushek, 1992). Others have also documented a dilution...
effect, whereby a larger number of siblings decreases parental resources (Blake, 1989) and less education investment per child (Psacharopoulos and Patrinos, 1997). Adolescent childbearing also has negative implications for the cognitive and educational development of the mother and child. Studies have shown that first-born children of adolescent mothers score lower on some cognitive/intelligence tests early in life, as well as have lower school attendance and achievement in reading and arithmetic (Hofferth, 1987; Karra and Lee, 2012).

Finally, distal drivers—like poverty or household income—negatively affect parents’ ability to invest in children’s schooling, including in school enrollment/other fees and buying textbooks. Government investment in education, with household income, may also indirectly affect student achievement through its impact on student-, teacher-, and school-level factors.

As previously noted, we could not test these evidence-based causal paths using SEM because of the low number of observations of the dependent variable. Instead, we used multiple regression to model this variable. Due to the lack of country-level data for student effects, teacher effects, and school effects, three proxies were included in the regression analysis: pupil-teacher ratio, quality of teacher training, and quality of primary education. We also tested the TFR, adolescent birth rate, public sector spending on education, mean years of female education of the cohort 25 years+, and GDP per capita. The final model included the quality of primary education and the adolescent birth rate. The response variable included the mean of values from 2011–2014. The quality of primary education variable included the mean of values from 2011–2015, while retrospective values from 2000 to 2005 were used for the adolescent birth rate to account for the primary school child’s age (approximately 13 years old). The adolescent birth rate and the quality of primary education explained 48 percent of the variation in the proportion of children at the end of primary school achieving at least a minimum proficiency level in reading (see Table 13).

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(proportion of children at the end of primary school achieving minimum proficiency in reading)</td>
<td>Adolescent birth rate</td>
<td>-.21</td>
<td>-3.12</td>
<td>.003</td>
<td>.48</td>
<td>51</td>
</tr>
<tr>
<td>Quality of primary education</td>
<td>5.03</td>
<td>2.06</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SDG indicator 6.1.1: Proportion of the population using safely managed drinking water services**

Safely managed drinking water services consist of improved water sources that meet three criteria: they are (1) located on premises (dwelling, yard, or plot), (2) available when needed, and (3) free from contamination. Improved water sources include piped water, boreholes, or tube wells; protected dug wells; protected springs; rainwater; and packaged or delivered water (UN Statistics Division, 2017).

Safely managing drinking water can be conceptualized as an availability, accessibility, and quality issue (WHO, 2017b). For availability, weak public sector management of water resources can challenge availability and, therefore, use—especially by those living beyond the reach of centralized systems or in informal urban settlements. Challenges can also include insufficient investment in the development of water systems, delivery mechanisms, maintenance and/or repairs (WHO, 2017b; Hunter et al., 2010); inflexibility in land regulation for residential use
Financial constraints prevent families from investing in household water systems, addressing needed repairs, or paying for recurrent costs like water tariff or user fees, bottled/vendor water, and/or maintenance fees (WHO and UNICEF, 2017). Governments in some low-income countries struggle to recover costs of improved drinking water where provided, leading to deterioration of those systems (Hunter et al., 2010). Moreover, hygiene education is needed so families know how and when to properly use services after they are available. Even when water is available and accessible, its quality may be degraded. Quality issues include microbial pollution because of fecal contamination, chemical pollution from industrial and household sources, industrial pollution through waste disposal, as well as agricultural pollution from pesticides (Ezenwaji et al., 2015; Grady et al., 2014).

Building on this evidence and available data, we tested 14 predictors on the outcome—predictors related to availability (annual freshwater withdrawals for domestic use, water productivity, population growth rate, TFR, public sector management, etc.), accessibility (GDP per capita), and quality (fertilizer consumption). The final model included GDP per capita and TFR, explaining 77 percent of the variation in the proportion of the population using safely managed drinking water (see Table 14).

### Table 14. Population Using Safely Managed Drinking Water Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of the population using safely managed drinking water services</td>
<td>Ln(GDP per capita)</td>
<td>10.80</td>
<td>8.85</td>
<td>.000</td>
<td>.77</td>
<td>84</td>
</tr>
<tr>
<td>TFR</td>
<td></td>
<td>-9.42</td>
<td>-5.70</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SDG indicator 6.2.1: Proportion of the population using a safely managed sanitation facility**

A safely managed sanitation facility hygienically separates excreta from human contact. These improved sanitation facilities should not be shared with other households and the excreta produced should be either (1) treated and disposed of in situ; (2) stored temporarily and then emptied, transported, and treated off-site; or (3) transported through a sewer with wastewater and then treated off-site. Examples of improved facilities include flush/pour flush to piped sewer systems, septic tanks or pit latrines, ventilated improved pit latrines, and composting toilets or pit latrines with slabs (WHO and UNICEF, 2017).

As with SDG indicator 6.1.1, using a safely managed sanitation facility can also be conceptualized as an availability, accessibility, and quality issue. Weak public sector management of sanitation services includes inadequate financial prioritization of sewage system setup or maintenance; poor coordination among national and sub-national actors, resulting in non-fulfillment of responsibilities; and insufficient micro-credit or lending opportunities for private sanitation setup (Davis, 2004; WHO, 2004).
Across all sources of funding, households contribute the largest share of resources to the funding of sanitation services (World Bank, 2017). Financial constraints can prevent households from investing in improved facilities independently, and unaffordable recurrent costs (wastewater tariffs, public toilet user fees, maintenance costs, etc.) may inhibit access to public sources (WHO and UNICEF, 2017). Sanitation services are also challenged by the demands resulting from population growth, placing a strain on governments already struggling to prioritize needs and sectoral investments. Quality is especially challenged by the improper disposal of wastewater; even when a sewage connection exists, poor treatment and disposal occur frequently in low-income countries, with consequences for human health and the environment (World Bank, 2017). In households with self-supply solutions, inadequate knowledge of proper use and disposal—because of inadequate hygiene education—also compromises health and well-being.

Building on this evidence and available data, we tested 14 predictors on the outcome—predictors related to institutions for sanitation services and availability of services (public sector management, property rights, TFR, population growth, urbanization, etc.) and accessibility (GDP per capita). The final model included only GDP per capita, explaining 53 percent of the variation in the proportion of the population using safely managed sanitation services (see Table 15).

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of the population using safely managed sanitation services</td>
<td>Ln(GDP per capita)</td>
<td>18.98</td>
<td>9.23</td>
<td>.000</td>
<td>.53</td>
<td>77</td>
</tr>
</tbody>
</table>

**SDG indicator 8.7.1: Proportion of children engaged in child labor**

Child labor is measured as the proportion of children ages 5–17 engaged in labor, divided by the total number of children (5–17 years old) in the population (UN Statistics Division, 2017). While the operational definition of child labor varies by data source and across countries, children are generally classified as laborers when they are either too young to perform certain activities or are involved in activities that negatively impact their health, personal development, or dignity. Generally, a child is considered to be involved in labor if they meet the following age-specific thresholds during a reference week: (a) children 5–11 years old who did at least one hour of economic activity; (b) children 12–14 years old who did at least 14 hours of economic activity; and (c) children 15–17 who did at least 43 hours of economic activity (UN Statistics Division, 2017). The extent to which work within the household constitutes child labor remains contested. In Multiple Indicator Cluster Surveys, age-specific thresholds for time spent on household chores are included in the child labor definition.¹²

Poverty is one of the main drivers of child labor (Webbink et al., 2013). Financial limitations may make child labor a necessity for survival and, therefore, a key supplement to parent income. According to this theory—termed the luxury axiom—a household would not require children to work if its income from non-child labor sources were sufficiently high (Basu and Van, 1998).

¹² (a) Children 5–11 and 12–14 years old: 28 hours or more and (b) children 15–17 years old: 43 hours or more.
Related to this, children from large families are more likely to work than those from smaller families because parent income/resources are diluted (ILO, 2004).

Moreover, parent education also impacts child labor—an uneducated parent may not recognize all the benefits of child education or the trade-off between work and attending school (Edmonds and Pavcnik, 2005). Likewise, parents working in agriculture or basic industry are more likely to see value in introducing children to their trade rather than having them pursue education. This is especially true for mothers working in agriculture, who are more likely to bring their child to work beside them, creating de facto conditions for a child to begin work (Webbink et al., 2013). Additionally, children are more likely to work if they come from families in which one or both parents are absent/deceased (Webbink et al., 2013).

While attending school and working are not necessarily incompatible, limited education opportunities—because of high financial costs, geographic access, and/or perceived or real quality issues like teacher absenteeism—might cause parents to choose work as the best use of a child’s time (Edmonds and Pavcnik, 2005; Webbink et al., 2013). Norms and values also play a role in whether children work or not. In contexts with stronger patriarchal values, parents are more likely to invest in the education of sons. Daughters, however, may be kept out of school and encouraged to work inside the home (Webbink et al., 2013). Similarly, the more empowered women are, the more capable they are of using their influence to benefit their children, and, therefore, support majority or exclusive schooling rather than housework or external work (Das and Mukherjee, 2007).

As noted in the Methodology section, we were unable to test these evidence-based causal paths using SEM because of the low number of observations of the dependent variable. Instead, we used multiple regression to model this variable. Ten independent variables were tested, including educational variables—female mean years of education for people ages 25+, quality of primary education, household funding of education, number of years of free education in the country, and the country’s expenditure in education per primary student—and other distal variables (e.g., share of employment within agriculture, TFR, income per person, women’s empowerment, and the prevalence of orphans). The best model was estimated with TFR, GDP per capita, and the number of years of free education (see Table 16). This model explained 52 percent of the variation in child labor.

### Table 16. Child Labor Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of children engaged in child labor</td>
<td>Years of free education</td>
<td>-0.73</td>
<td>-1.79</td>
<td>.076</td>
<td>.52</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Ln(GDP per capita)</td>
<td>-3.89</td>
<td>-3.70</td>
<td>&lt; .001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total fertility rate</td>
<td>2.41</td>
<td>3.20</td>
<td>.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SDG indicator 11.1.1: Proportion of the urban population living in slums**

This SDG indicator was originally conceptualized as “the proportion of urban population living in slums, informal settlements or inadequate housing” (United Nations, 2018, pg. 11). Because of data limitations, this indicator is currently measured as “the proportion of the urban population living in slums.” A slum household is one in which the inhabitants experience one or more of the following deprivations: the lack of (1) access to an improved water source, (2) access to improved sanitation facilities, (3) sufficient living area, (4) housing durability, or (5) security of tenure (UN Statistics Division, 2017). Beyond the most immediate drivers of slum dwelling—
operationalized in the definition of the indicator—three groups of factors affect housing: availability, accessibility, and quality.

Governance and management of the housing sector directly impacts the growth of slums as it affects the supply, quality, and accessibility (affordability) of homes. Escalating land cost price low-income households out of the market (UN Habitat, 2016). Inappropriate regulatory frameworks cause unequal and inefficient land development; where regulations have been relaxed to encourage residential construction and small firm entry, benefits have accrued for the lowest income groups (UN Habitat, 2016). Despite promising intentions, housing subsidies have not significantly curtailed slum living, given a built-in bias against the poorest households (e.g., minimum income threshold and proof of formal employment) (UN Habitat, 2016), while slum upgrading programs have proven more effective (UN Habitat, 2016; Arimah, 2010).

Inadequate income also curtails household access to adequate living conditions. Low-income households are more likely to reside in slums, especially if employed in the informal sector. As individuals earn more money, they are better able to demand and realize improved housing conditions. Increased income within the formal sector means increased tax revenue, funds that can be used for slum upgrading programs (Arimah, 2010).

Rapid population growth may amplify urban problems by straining government systems (Arimah, 2010; Roy et al., 2014). Rapid urbanization places a direct strain on city authorities to provide adequate housing and infrastructure, as well as to effectively manage the process and consequences of urban development (UN Habitat, 2016). The natural growth rate of urban communities—or the urban growth rate—exerts the same effect, testing the capacities of the government (Ooi and Phua, 2007). All three demographic factors may push individuals to live in slum or squatter settlements because of the limited availability of adequate accommodations.

Building on this evidence and available data, we tested 12 predictors linked to governance of the housing sector (business regulatory environment, public administration, etc.), accessibility (GDP per capita and social protection programs), and demographics (urban growth rate, urbanization, etc.) on the outcome. We did not test property rights because of its repetitiveness with the measurement of the outcome indicator (lack of security of tenure). The final model included the variables GDP per capita and the growth of the urban population. This final model explained 51 percent of the variation in the proportion of the urban population living in slums (see Table 17).

Table 17. Proportion of Urban Population Living in Slums Estimation

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Predictor Variables</th>
<th>Coefficient</th>
<th>T Statistic</th>
<th>P-Value</th>
<th>R²</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of the urban population living in slums</td>
<td>Ln(GDP per capita)</td>
<td>-12.65</td>
<td>-5.69</td>
<td>0.00</td>
<td>.51</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Growth rate of the</td>
<td>3.72</td>
<td>2.04</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>urban population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex A. Data Definitions and Sources

Data from 231 countries and territories were used to build the FP-SDGs Model, although not all countries had complete data for each variable. Key variables used to build the model and the data source for each are defined below. These variables were grouped based on policy area addressed. When the user applies the FP-SDGs Model, data for all required variables are auto-populated from the model’s DATABASE sheet, and the default source and date for each is clearly labeled in the BASE YEAR DATA INPUTS worksheet. The user can enter alternate data sources for the baseline and policy variables, if desired. The variables in Table 18 are found in the BASE YEAR DATA INPUTS sheet, and a subset are also included in the SET POLICY GOALS sheet.

Table 18. Key Variable Definitions and Data Sources

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>YEARS</th>
</tr>
</thead>
</table>
### Demographic Data

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>YEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Postpartum insusceptibility</strong></td>
<td>Number of months after childbirth when half of women are no longer protected against pregnancy either by postpartum amenorrhea or abstinence from sex.</td>
<td>ICF International. 2016. Demographic and Health Surveys [Datasets]. Calverton, Maryland: ICF International.</td>
<td>Latest available data by country between 2005 and 2015</td>
</tr>
<tr>
<td><strong>Total fertility rate</strong></td>
<td>The average number of children a hypothetical cohort of women would have at the end of their reproductive period if they were subject during their whole lives to the fertility rates of a given period and if they were not subject to mortality. It is expressed as children per woman.</td>
<td>United Nations, Department of Economic and Social Affairs, Population Division. 2015. <em>World Population Prospects: The 2015 Revision</em>, DVD Edition.</td>
<td>2010–2015 data for available countries</td>
</tr>
</tbody>
</table>
### Family Planning-Sustainable Development Goals Model Methodology and User’s Manual

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>YEAR</th>
</tr>
</thead>
</table>

### Health Data

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>YEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent births at risk</td>
<td>Percentage of children born in the last five years in any risk category. Risk categories include births to women under 18 years of age or over 34 years of age, births of order 4 or higher, and births within 24 months of a previous birth.</td>
<td>ICF International. 2016. Demographic and Health Surveys [Datasets]. Calverton, Maryland: ICF International.</td>
<td>Latest available data by country between 2005 and 2015</td>
</tr>
<tr>
<td>Maternal mortality ratio</td>
<td>Maternal deaths per 100,000 live births. A maternal death is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes.</td>
<td>United Nations. “SDG Indicators Global Database.” Indicator 3.1.1. Retrieved July 2016 from <a href="http://unstats.un.org/sdgs/indicators/database?indicator=3.1.1">http://unstats.un.org/sdgs/indicators/database?indicator=3.1.1</a>.</td>
<td>2015 data for available countries</td>
</tr>
<tr>
<td>VARIABLE</td>
<td>DESCRIPTION</td>
<td>SOURCE</td>
<td>YEAR</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
</tr>
<tr>
<td>Proportion of the population using safely managed drinking water services</td>
<td>Proportion of the population using safely managed drinking water services, or those that consist of improved water sources that meet three criteria: 1) they are located on premises (dwelling, yard, or plot), 2) available when needed, 3) and free from contamination. Improved water sources include piped water, boreholes or tube wells, protected dug wells, protected springs, rainwater, and packaged or delivered water.</td>
<td>United Nations. “SDG Indicators Global Database.” Indicator 6.1.1. Retrieved November 2017 from <a href="http://unstats.un.org/sdgs/indicators/database?indicator=6.1.1">http://unstats.un.org/sdgs/indicators/database?indicator=6.1.1</a>.</td>
<td>2011–2015 data for available countries</td>
</tr>
<tr>
<td>Proportion of the urban population living in slums</td>
<td>Proportion of urban population living in slums. A slum household is one in which the inhabitants suffer one or more of the following deprivations: lack of (1) access to an improved water source, (2) access to improved sanitation facilities, (3) sufficient living area, (4) housing durability, or 5) security of tenure (UN Statistics Division, 2017).</td>
<td>United Nations. “SDG Indicators Global Database.” Indicator 11.1.1. Retrieved July 2016 from <a href="http://unstats.un.org/sdgs/indicators/database?indicator=11.1.1">http://unstats.un.org/sdgs/indicators/database?indicator=11.1.1</a>.</td>
<td>2009 data for available countries</td>
</tr>
</tbody>
</table>
## Proportion of population using safely managed sanitation services

Proportion of the population using safely managed sanitation facility, or one that is designed to hygienically separate excreta from human contact. These improved sanitation facilities should not be shared with other households, and the excreta produced should be either: 1) treated and disposed of in situ; 2) stored temporarily and then emptied, transported and treated off-site; or 3) transported through a sewer with wastewater and then treated off-site.


## Minimum dietary diversity

Percentage of children 6–23 months of age who ate from at least 4 (out of 7) pre-defined food groups during the previous day.


## Skilled birth attendance

Percentage of deliveries attended by personnel trained to give the necessary supervision, care, and advice to women during pregnancy, labor, and the postpartum period; to conduct deliveries on their own; and to care for newborns.


## Antenatal care

Percentage of women aged 15–49 years attended at least four times during pregnancy by any provider.


## Tetanus immunization

The percentage of births by women of child-bearing age who are immunized against tetanus.


### Education Data

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>YEAR</th>
</tr>
</thead>
</table>
## Family Planning-Sustainable Development Goals Model Methodology and User’s Manual

### Economic Data

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>YEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial employment</strong></td>
<td>Population (age 15+) in employment. Employment is defined as persons of working age who, during a short reference period, were engaged in any activity to produce goods or provide services for pay or profit, whether at work during the reference period (i.e., worked in a job for at least one hour) or not at work due to temporary absence from a job, or to working-time arrangements.</td>
<td>International Labour Organization. ILOSTAT. Retrieved July 2016 from <a href="http://www.ilo.org/global/statistics-and-databases/lang-en/index.htm">http://www.ilo.org/global/statistics-and-databases/lang-en/index.htm</a>.</td>
<td>2016 data for available countries</td>
</tr>
<tr>
<td>VARIABLE</td>
<td>DESCRIPTION</td>
<td>SOURCE</td>
<td>YEAR</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td><strong>Labor force participation rate</strong></td>
<td>Proportion of a country's working-age population (ages 15–64) that engages actively in the labor market, either by working or looking for work. The labor force participation rate is calculated by expressing the number of persons in the labor force as a percentage of the working-age population. The labor force is the sum of the number of persons employed and the number of persons unemployed (without a job but available for work).</td>
<td>International Labour Organization. ILOSTAT. Retrieved July 2016 from <a href="http://www.ilo.org/global/statistics-and-databases/lang-en/index.htm">http://www.ilo.org/global/statistics-and-databases/lang-en/index.htm</a>.</td>
<td>2016 data for available countries</td>
</tr>
<tr>
<td>VARIABLE</td>
<td>DESCRIPTION</td>
<td>SOURCE</td>
<td>YEAR</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>GDP per capita is the GDP divided by the population size. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without deducting for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2010 USD currency.</td>
<td>World Bank. “Indicators: GDP Per Capita, 2010 Constant $US.” Retrieved July 2016 from <a href="http://data.worldbank.org/indicator">http://data.worldbank.org/indicator</a></td>
<td>Latest available data by country 2010–2015</td>
</tr>
<tr>
<td>VARIABLE</td>
<td>DESCRIPTION</td>
<td>SOURCE</td>
<td>YEAR</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Capital stock growth rate</td>
<td>Percent change year-on-year in capital stock.</td>
<td>No default data available; use domestic sources to estimate.</td>
<td>—</td>
</tr>
<tr>
<td>Capital stock depreciation rate</td>
<td>Rate of gradual decrease in the economic value of capital stock in an economy.</td>
<td>No default data available; use domestic sources to estimate.</td>
<td>—</td>
</tr>
</tbody>
</table>
Annex B. Structural Equation Modeling Path Diagrams

Figure 18. Prevalence of Stunting

Note: * p ≤ 0.05; ** p < 0.01; *** p < 0.001. Numbers in diagram show coefficients. Covariances between exogenous variables are established as a standard procedure when using the maximum likelihood for missing values method. They are not presented here because they do not alter the coefficients for predicting the outcome.
Figure 19. Maternal Mortality Ratio

Total fertility rate

70.31***

Antenatal care

0.28***

Female education (mean years)

2.44***

GDP per capita (Ln)

2.88**

Skilled birth attendance

-3.48***

Adolescent birth rate

1.32**

Maternal mortality ratio

Note: * p ≤ 0.05; ** p < 0.01; *** p < 0.001. Numbers in diagram show coefficients. Co-variances between exogenous variables are established as a standard procedure when using the maximum likelihood for missing values method. They are not presented here because they do not alter the coefficients for predicting the outcome.
Figure 20. Under-Five Mortality Rate

| Variable                        | Coefficient | p-value  
|--------------------------------|-------------|----------
| Total fertility rate           | 9.55***     |          
| Antenatal care                 | 0.28***     |          
| Female education (mean years)  | 2.53***     |          
| GDP per capita (Ln)            | 2.67*       |          
| Skilled birth attendance       | -0.35**     |          
| Adolescent birth rate          | -1.54*      |          
| Tetanus immunization           | 0.19***     |          
| Under-five mortality rate      | -0.05*      |          
|                                | -0.42*      |          

Note: * p ≤ 0.05; ** p < 0.01; *** p < 0.001. Numbers in diagram show coefficients.

Co-variances between exogenous variables are established as a standard procedure when using the maximum likelihood for missing values method. They are not presented here because they do not alter the coefficients for predicting the outcome.
References


Marks, G. 1998. “Attitudes to School Life: Their Influences and Their Effects on Achievement and Leaving School.” LSAY Research Reports n.5.


For more information, contact:

Health Policy Plus
Palladium
1331 Pennsylvania Ave NW, Suite 600
Washington, DC 20004
Tel: (202) 775-9680
Fax: (202) 775-9694
Email: policyinfo@thepalladiumgroup.com
www.healthpolicyplus.com