Supporting Information (SI) Appendix

for

Education Under SSP Manuscript(Title not yet finalized)

Table of Contents

This document outlines our approach to projecting educational schooling and adult attainment using the International Futures (IFs) model, an integrated assessment tool covering social, economic, and environmental variables across 188 countries. The IFs education model forecasts enrollment, graduation rates, and attainment levels, deriving global and regional estimates from population-weighted country projections.

We begin with an overview of the IFs modeling system, followed by a breakdown of the education model's key components. We then describe the historical data used for the model, followed by detailed sections on its main components: schooling flows, financial flows, and educational attainment. Next, we outline our scenario development process and the underlying logic. Finally, we provide instructions for replicating the model and scenarios. The IFs Model Wiki (1) offers a more extensive description of the modeling methodology, model coverage, and data sources, along with illustrative diagrams.

1. International Futures Model

International Futures (IFs) is a freely available integrated modeling system designed for long-term global, regional, and national projections across demographics, economics, education, energy, environment, health, infrastructure, and governance. It's hard-linked, discipline-specific sub-models advance recursively in annual time steps through 2100.

These interconnected sub-models (Figure S1) influence and respond to each other through analytical equations and algorithms based on historical data, capturing structural, accounting, and flow dynamics. While most interactions occur within national boundaries, global processes like trade and migration are handled by default through a pooled approach, but bilaterally as an option.

IFs produces a comprehensive "base case" projection, a data-driven reference scenario that forecasts long-term trends based on historical patterns, current policies, and structural relationships. This serves as a reference for exploring alternative scenarios and policy interventions, allowing users to modify parameters and assess potential impacts.

Relative to other Integrated Assessment Models (IAMs), IFs stands out for its broader issue coverage, user-friendly interface, and policy adaptability. Its strength lies in evaluating the combined effects of policy choices, capturing dynamic feedback loops across its component models.

For example, the economic and demographic sub-models are tightly linked to the education model, reflecting how educational investments influence economic growth and demographic trends and how they in turn affect education. This integrated approach allows IFs to analyze effects of policy decisions on schooling, workforce productivity, and long-term development.

2. IFs Education Model

An integrated modeling approach requires a comprehensive representation of the system and a unified analytical framework to capture dynamic interactions that shape and are shaped by the system's key outcomes. However, most global education models (2–6), whether they rely on statistical or structural projection methods, have relatively limited interactions with other socio-economic factors. In contrast, IFs embeds education within a larger system of interconnected demographic, economic, and governance factors, ensuring dynamic interactions and feedback loops. Two of its key strengths lie in endogenizing the interactions between education and broader society and using a full stock-and-flow system that tracks students from entry into and progression through school to their long-term educational attainment in adulthood and the workforce.

IFs recognizes that education demand and supply are shaped by economic growth, labor market needs, and public finance constraints. A growing economy increases demand for skilled workers, reinforcing education investment, while government budget limitations, especially in lower-income countries, restrict early-stage education expansion. Education costs—primarily teacher compensations (7)—are modeled to increase with per capita income, reflecting rising labor opportunity costs. Enrollment and funding demands are linked to age-specific population projections from the IFs cohort-component population model. At the same time, IFs captures how education feeds back into the economy, as a more educated workforce drives productivity, boosts growth and government revenue, and enables further investment in education. These two-way relationships (Figure S2) ensure that education is not analyzed in isolation but as part of a dynamic, evolving system that shapes and is shaped by broader long-term development trajectories.

Another key feature of the IFs education model is its integrated stock-and-flow accounting system. It simultaneously tracks student movement through the schooling system (flows) and adult educational attainment (stocks). Unlike models that statistically estimate adult educational attainment, IFs explicitly models how students enter, progress, graduate, or drop out, ensuring a direct link between schooling experiences and long-term educational outcomes. The model dynamically projects intake, survival, enrollment, and completion rates, ensuring that cohort movements align with educational progression from elementary entry through adulthood. By linking student flows to adult educational attainment and integrating attainment with the broader economy and demography, IFs provides a structurally grounded and policy-sensitive projection of educational attainment. Its ability to capture the interconnected nature of education, economic development, and demographic change makes the IFs education model a powerful tool for long-term planning. Further details on the data used to initialize the model, along with the key dynamics of the schooling flow and attainment model, are provided in the following sections.

3. Data Used in the IFs Education Model

The structure of the education model relies critically on data availability, the imputation of missing values, and the reconciliation of data from multiple series, whether from the same or different sources. IFs prepares data through a dedicated "preprocessor," which constructs base-year values—currently for 2020. This preprocessor reads from the IFs historical database, where variables are organized in a country-year format, and applies imputation methods and consistency checks to generate a complete dataset for all 188 modeled countries. Once this initialization is complete, the model can be run to produce the Base Case scenario. The IFs interface also supports the specification and analysis of alternative scenarios, all anchored in the same 2020 base-year values. The education model uses historical data to initialize three key variable categories: student flow rates (e.g., entrance or graduation), financial data (e.g., costs, spending), and adult educational attainment. This section describes the data sources and preprocessing steps; Sections 4 and 5 explain how the data feed into the student flow and financing models, while Section 6 addresses the progression of educational attainment into adulthood.

Data Sources

The education model data are sourced from recognized repositories of global education statistics. Student and financial flow data are primarily obtained from UNESCO Institute for Statistics (UIS) (8), while educational attainment data come from the Wittgenstein Centre (WIC) (9, 10). Table S1 summarizes the data sources for all education variables as well as other variables used in initializing the education module of the IFs model.

To ensure comparability across countries, the UNESCO Institute for Statistics (UIS) follows the International Standard Classification of Education's 2011 framework, ISCED 2011 (11). The IFs education model incorporates four education levels, consistently organized for all countries in the UIS database: Primary (ISCED level 1), Lower Secondary (ISCED level 2), Upper Secondary (ISCED level 3), and Tertiary (first-degree programs of ISCED levels 5, 6 or 7). The UIS database provides country-specific values for the age of entry and the duration of each education level up through upper secondary, which IFs implements. To simplify variation in program length both within and across countries, IFs represents the full duration of tertiary education as five years. This age mapping is essential for computing schooling flow rates, many of which are defined as a share of the relevant age-group population. For example, enrollment at a given level is calculated as the number of children enrolled expressed as a percentage of the total population at the appropriate schooling age, from entry to the final grade.

The UIS data acknowledge that, in practice, children may start school late, repeat grades, or take breaks. To account for misalignment between age and grade/schooling level, UIS provides two types of flow rates where applicable.

- Gross rates include all pupils with the defined schooling attribute, even those older or younger than the expected age, divided by the population at the expected age, and can exceed 100%.

- Net rates, in contrast, count only those students, entrants, or graduates who are of the expected age for a given level, divided by the population at the expected age.

IFs projects both net and gross rates at the elementary level to estimate the large number of school-aged children who remain out of school in many developing countries. Over time, gross rates tend to converge toward net rates as of-age enrollment rates rise.

For all other schooling levels, IFs projects only gross rates, as they better reflect system capacity and demand. Not all student flow rates follow this gross-net distinction. For example, the survival rate, which is defined as the share of entrants who reach a certain grade, is inherently limited to 100%.

The grade distribution of students is initialized in two ways: either using grade-specific student flow data, or—when such data are unavailable, unreliable, or have limited coverage—using more aggregate, level-wide flow rates combined with assumptions about their distribution across grades. For primary and lower secondary levels, grade-specific enrollment counts from UIS are combined with single-year population estimates from the UN Population Division's (UNPD) World Population Prospect (12) to calculate enrollment rates by grade. UIS also provides grade-specific dropout rates for these levels, which the IFs education model incorporates.

For upper secondary and tertiary levels, entrance, progression, and exit flow rates are used to compute the initial-year grade distribution through a distribution algorithm. This algorithm makes a simplifying assumption: that current students progressed through the level with current entrance and graduation rates, and that dropouts were evenly distributed across the intervening grades.

UIS (8) compiles and publishes financial flow data on education spending per student and total funding at each level, as well as for the education sector as a whole. These data are sourced primarily from UIS, but also from the World Bank's World Development Indicators (WDI) (13) and Education Statistics (EdStats) (14), both of which derive their data from UIS sources (8). The GDP data in IFs come from the World Economic Outlook database (15).

While IFs draws on both the Wittgenstein Centre for Demography and Global Human Capital (WIC) and Barro-Lee (B-L) (16) data for educational attainment, this paper uses only WIC data—resorting to Barro-Lee data solely to fill in values for countries missing from the WIC dataset (9, 10).

Data Preprocessing in IFs Model

While UIS provides good data coverage, gaps remain in both student flow and financial data. In contrast, the attainment data has nearly complete geographic coverage. To ensure base-year data are available for all variables and all countries, IFs uses a systematic imputation process to fill missing values. The same methods are applied when initializing a given variable, though the set of imputation techniques may vary based on disaggregation level (e.g., grade vs. schooling level), data availability, reliability, country coverage, and time span. The techniques include closest-year substitution, cross-sectional estimation, longitudinal extrapolation, or full-flow reconstruction with available data. The model pre-processor ensures internal consistency between observed and imputed data, and across student flows at the same or contiguous levels for each country. These data imputation and reconciliation methods are described below.

The education pre-processor begins initialization by reading base-year values for each variable for all countries from the historical IFs database, which compiles data from the sources listed in Table 1. When a data value for the base year (2020 in IFs 8.26) is missing, the imputation first attempts to use the value from the closest available year within a defined window, typically five years from the base year.

If no recent data exist, missing values are estimated in most cases using cross-sectional relationships typically derived from the most recent data available for all countries, with the variable being initialized regressed against an indicator of the country's level of development and educational demand, most often GDP per capita at purchasing power parity (GDPPCPPP). GDPPCPPP is an essential IFs data series with full country coverage for the base year and a sufficiently long preceding period. Depending on the variable, the imputed value is derived by using either the level of GDPPCPPP in the base year or the change in GDPPCPPP between the base year and the nearest available year for which the imputed series has a value in the country being filled in.

It is possible to construct a grade-specific distribution of students for the current year using available data on entrance, dropout, and enrollment rates for the entire level—provided these rates are assumed to be the same across all cohorts currently in school. Building on this assumption, the full-flow reconstruction method uses available values for these rates to estimate a missing flow rate when others are known. The method is also used to ensure internal consistency when all flow rates are available—either from data or other imputation techniques.

When inconsistencies arise—particularly at the upper secondary and tertiary levels where grade-specific data are not available—values derived from historical data are prioritized over imputed values. Among available rates, conceptually clearer indicators such as enrollment are generally preferred over others, such as survival; similarly, intake rates are favored over survival rates.

For primary and lower secondary levels, where grade-specific data are available, the flow reconstruction method is used to derive level-specific rates when only one of the flow-rate values is missing. When multiple values are missing, cross-sectional estimation is first used to impute all but one, and the reconstruction method is then applied to estimate the final missing value. These level-specific flow rates are subsequently reconciled with the grade-level data initialization.

For primary and lower secondary levels, grade-level enrollment data is available as headcounts, which we use in conjunction with single-year population data to compute grade-specific enrollment rates. If data for only one grade is missing in a country, it is estimated using available data for all other grades and the previously filled total enrollment for that level. If multiple grades are missing, different estimation techniques are applied: the last grade is estimated using completion and survival rates, while the first grade is estimated using entrance or transition rates. For other missing grades, a linear trend based on available grade-level data is used to interpolate values. If no grade-specific data exists, flow-rates for all grades are initialized algorithmically using the same method described earlier in this section to compute grade distributions from level-wide flow rates and level durations.

For grade-specific dropout rates, we use regression functions driven by GDP per capita. Additionally, dropout rates for all grades must be consistent with the survival rate to the last grade (or the graduation rate for lower secondary), which allows us to derive the dropout rate when only one grade is missing. The survival rate and grade-specific dropout rates also serve as a consistency check for estimated dropout values.

Since both grade-specific and level-specific data come from the same source, they are usually aligned. However, reconciliation is sometimes needed for level-specific flow rates involving a single grade or single-year age group—such as entrance or completion rates—which may occasionally misalign with their grade-specific equivalents, like enrollment in the first and last grades. In such cases, we check whether the grade-specific or level-specific data has been imputed and prioritize reported data over imputed values. If both are imputed, we retain the grade-specific imputation. We follow a similar approach when reconciling the sum of grade-specific dropout rates with the survival or graduation rate for the level. In some cases, when there is a significant discrepancy between the two values being reconciled, we take the midpoint as a balanced estimate.

WIC and B-L estimate attainment series ((9, 10, 16) using population and education data, but these series are typically only reported up to about 5 to 10 years before the regularly advancing base year of IFs, most recently 2020. Given that educational attainment grows slowly and follows a predictable pattern, we extend them to the base year using longitudinal extrapolation for each country's trajectory.

Per-student public costs at each of the four education levels are reported as shares of per capita income (GDPPCPP) by UIS(8), facilitating comparisons of national commitments to education and learning quality. IFs uses cross-sectional regressions, based on the most recent data and driven by per capita income (GDPPCPP), to estimate missing values.

Once we construct a complete set of initial data points for all countries and variables, the model is ready to generate forecasts using dynamic projections. In the next sections, we describe these dynamic behaviors for student flows and education financing and educational attainment.

4. Student Flow Model

The IFs education model projects enrollment, graduation, and other schooling flow outcomes by gender for each country across four education levels: primary, lower secondary, upper secondary, and tertiary. The fundamental unit in the schooling model is a grade, with each level comprising multiple grades. The model tracks students as they enter primary school and progress through grades, where they can either advance, repeat, or drop out due to personal circumstances or systemic failures.

The model simulates student flows grade by grade, initializing each grade with grade-specific enrollment and dropout rates—using historical data for primary to lower secondary levels, and distributions derived from student-flow reconstruction for other levels, as described in an earlier section. Each year, grade-level enrollment is updated by carrying forward students from the previous grade, adjusting for dropouts and repeaters. The first grade at each level accounts for new entrants, either through fresh intake (first entry into the system) or transition from the previous level, while the last grade reflects graduation or completion rates. Student flow rates, when multiplied by the relevant age-group population, yield student counts.

The key dynamics of the student flow rates are determined using analytical functions, typically derived by regressing the most recent flow rate data against common development indicators such as real GDP per capita at PPP. As noted elsewhere in the SI, the flow rates they produce are subsequently adjusted, as needed, based on education expenditure. These functions are estimated separately by gender to account for differences in educational progress between girls and boys and are also used to impute missing data for most student flow variables. Figure S3 illustrates two such functions—net intake rates in primary education for boys and girls—plotted against GDP per capita in 1000 PPP (2017) dollars.

Flow rates are used to compute future grade-level enrollments, which are initialized in the base year using either reported data or the reconstructed flow methods described earlier. These enrollments evolve over time through a flow accounting process. Below, we illustrate this using the equation for primary grade-level enrollment (*pristudents*), where the subscripts *g*, *r*, *t*, and *p* denote grade, country, time period, and gender, respectively. Inflows include new entrants and students promoted from the previous grade ($pristudents_{g-1,\, t-1}$), while outflows account for promotions to the next grade, dropouts (EdPriDropoutByGrades), or graduations. Students who repeat (RptGr) a grade remain in the enrollment stock for that grade in the following year.

$$pristudents_{g,p,r,t} = pristudents_{g-1,p,r,t-1} * \left(1 - EdPriDropoutByGrade_{g-1,p,r,t} - RptGr_{g-1,p,r,t}\right) + pristudents_{g,p,r,t-1} * RptGr_{g,p,r,t} \qquad (1)$$

$$EDPRIENRG_{g,p,r,t} = \sum_{g=1}^{edprilen_r} pristudents_{g,p,r,t} \qquad (2)$$

These grade-level dynamics are computed year by year, and total enrollment for each education level (e.g., *PRIENRG* for primary) is derived by summing enrollments across all relevant grades. Student outflows—through graduation or dropout—are used to update the educational attainment of the appropriate age group. Figure 4 presents a representative schematic of the integrated student flows and the attainment stocks they update, which we describe in a later section.

IFs scenario parameters provide flexibility to simulate varying rates of student progression, allowing for faster or slower educational advancement depending on policy interventions, economic conditions, or other external factors. These variations ultimately shape enrollment patterns over time, as illustrated in Figure S5 for the SSP3 scenario used in our paper. The total enrollment serves as a key model output, informing various projections, including the demand for educational budgets, which we discuss in the next section.

5. Financial Flow Model

The IFs education model forecasts public funding of education using a demand-supply balancing approach, where demand is computed from total enrollment and per-student expenditures, while supply is determined by government budget constraints. Financial reconciliation is explained below. Any surplus or shortfall in public funding, relative to demand, affects student flows. Model users can simulate funding interventions using scenario parameters. The model does not currently include household spending on education.

Education is labor-intensive. Most education spending goes toward teacher salaries and other personnel compensation rather than capital or materials (17). To capture this relationship, IFs projects public spending per student at each education level using regression functions driven by income per capita

(GDPPCPPP), reflecting the tendency for education costs to rise with national income as the opportunity cost of labor increases. Country-specific variations in per student spending are accounted for by computing the initial-year difference between actual data and estimates, which gradually converges to zero as per-student expenditures align with global patterns.

Projected enrollment demand—derived from flow-rate dynamics—is multiplied by income-driven per-student cost projections for each level of schooling and summed to estimate unconstrained funding demand in each forecast year. This estimated demand is then subjected to overall funding constraints during the budget allocation process within the IFs sociopolitical model.

In the budgeting process, public spending on education and other categories—such as health, infrastructure, R&D, and defense—is estimated based on observed expenditure patterns for each category relative to economic development. For sectors where demand-side projections of expenditures are available (education, health, and infrastructure), user-adjustable priority parameters allow modification of the emphasis placed on current demand versus historically driven spending patterns. An additional user-defined multiplier, applicable to each spending category, provides further flexibility in adjusting expenditure levels. Total available government consumption—calculated in the economic model by balancing revenue and public spending—is then allocated across all categories. This allocation process normalizes projected sectoral spending, after reserving shares based on the specified priority parameters.

The allocated education budget is then compared to projected funding demand. If there is a surplus, it is allocated across levels based on the distance of each level from full enrollment, whereas in the case of a deficit, the shortfall is applied uniformly across all education levels. The budget impact—measured as the ratio of allocated to required funds—influences previously projected student flows. A funding surplus enhances student flow rates related to entrance (intake or transition from the earlier level) and progression (persistence/survival or graduation rate) non-linearly, with the greatest boost occurring when flow rates are near the midpoints of their possible range of value. In contrast, funding deficits have a more linear impact. Adjusted entrance, progression and exit flow rates are used to recompute enrollment projections.

Model users can alter budgetary allocations either by directly modifying budgetary parameters within scenarios or through adjusting flow rates, which in turn affect enrollment and ultimately impact the budget. We took the latter approach. Figure S6 presents budget expenditures as a percentage of GDP for low-income countries as a result of changes in student flows across the three SSP scenarios. These financing and student flow patterns ultimately influence adult educational attainment, which we discuss next.

6. Educational Attainment Model

This paper uses mean years of schooling (MYS) among the population aged 25 and above as the measure of adult educational attainment. In IFs, educational attainment is projected for all individuals aged 15 and above, disaggregated into five-year age-sex cohorts. For a given population group, MYS is computed by summing the average years of education obtained from both completed schooling and partial completion. At each time period (*t*) for a country (*r*), the average years of education for the population aged 15 and above (EDYRSAG15), disaggregated by gender (*p*), is calculated as the sum of four components: average or mean years from completed primary education (AvgYearsPriEdPop), completed upper secondary education (AvgYearsSecEdPop), completed tertiary education (AvgYearsTerEdPop), and partial completion (PartialYearsEdPop) by individuals who dropped out before graduating from any of these levels.

$$EDYRSAG15_{p,r,t} = AvgYearsPriEdPop_{p,r,t} + AvgYearsSecEdPop_{p,r,t} + AvgYearsTerEdPop_{p,r,t} + PartialYearsEdPop_{p,r,t} \tag{3}$$

Historical data from the Wittgenstein Centre (WIC) initialize completion rates for each five-year age-sex cohort. Completion rate projections from the IFs schooling flow model (e.g., EDPRICR for primary education) update these rates annually (e.g., EdPriPopPer for primary) for the appropriate age groups—primarily those aged 15–19 (cohort index c = 4) and 20–24—depending on the level of education.

$$EDPriPopPer_{c=4,p,r,t} = \left(4 * EDPriPopPer_{c=4,p,r,t-1} + EDPRICR_{r,p,t}\right)/5 \tag{4}$$

For older five-year age groups, completion rates for various levels of schooling evolve as individuals age. Each year, individuals at the upper end of their five-year age group move to the next cohort, bringing their educational attainment with them. This attainment is subtracted from the originating group and added to the receiving one, and new completion rates are calculated for each age group and each level of education, as shown below for primary. For simplicity, the model assumes that individuals do not acquire additional formal education later in life.

$$EdPriPopPer_{c=5\ to\ 21,p,r,t} = \left(EdPriPopPer_{c-1,p,r,t-1} + 4 * EdPriPopPer_{c,p,r,t-1}\right)/5 \tag{5}$$

Multiplying age-sex-specific completion rates by the number of individuals in each five-year group (*agedst*) and the duration of the corresponding education level (*edprilen*) yields the total number of education years from completion. This total is then divided by the population to obtain the average years of education for all adults. This computation is carried out separately for completed primary, secondary, and tertiary education—as illustrated in the equation below for primary.

.

$$AvgYearsPriEdPop_{p,r,t} = \frac{\sum_{c=4}^{NCohorts} \frac{EDPriPopPer_{c,p,r,t}}{100} * \boldsymbol{edprilen}_r * AGEDST_{c,p,r,t}}{\sum_{c=4}^{NCohorts} AgeDst_{c,p,r,t}} \tag{6}$$

The population projections in IFs are done in the demographic component of the integrated model, which accounts for demographic flows such as births, deaths, and migration—as well as the impact of education on these dynamics, as described in the IFs demographic model. For this paper, we have overridden International Futures' projection of total population with those from the Shared Socioeconomic Pathways (SSP) database(18).

Attainment years from partial completion of schooling at various levels are not reported separately in historical datasets. IFs initializes these values using base-year dropout rate data along with several simplifying assumptions. First, starting from the second grade (d = 2), the number of dropouts—calculated using the grade-specific enrollment rate (*Gr_Students*), dropout rate (*DropOutRate*), and relevant single-year population (*fagedst*) by age (c) and sex (p)—is multiplied by the number of years completed before dropping out. This process is repeated for all grades (d = *edprilen*) to obtain partial completion-related years for each level of schooling (e.g., *PartialPriPersYearsNew* for primary). The results are then summed to estimate partial completion for five-year adult cohorts close to school graduation ages (i.e., 15–19 and 20–24).

$$PartialPriPersYearsNew_{p,r,t} = \sum_{d=2}^{\boldsymbol{edprilen}_r} Gr_{Students_{d,p,r,t}} * DropoutRate_{d,p,r,t} * (d-1) * fagedst_{c,p,r,t} \tag{7}$$

A ratio of partial attainment to total attainment from primary completion among the 15–19 age group is then used to initialize partial attainment for older five-year age-sex cohorts. For projecting partial attainment in future years, the same dynamics are used as for completed attainment: current dropout rates are used to update partial attainment for graduation-age cohorts, and a cohort propagation method is applied to compute partial attainment for older age groups. Attainment from partial completions across all education levels and age groups (PartialYearsEdPop) is added to the MYS calculation, as mentioned earlier.

IFs derivation of MYS in the initial year, from various historical data and derivations, can be slightly different from the reported historical data on MYS. IFs computes this initial difference and adds it to the computed MYS.

7. Scenario Development

For this paper, we selected SSP2 (middle-of-the-road) and two SSPs (SSP3 and SSP5) that span a range of assumptions about the prospects for education. According to the SSP scenario narratives, education investments progress slowly in SSP2, decline in SSP3, and are strong in SSP5, as described in pages 172-175 of ref. (19). SSP1, which assumes low mitigation and adaptation challenges, suggests an educational investment trajectory similar to SSP5. SSP4's educational outcomes vary by fertility and income grouping (p. 176 of ref. (19)), potentially requiring more country-specific scenario construction than the global approach we employed. Given these considerations, we ultimately focused on SSP2, SSP3, and SSP5 to frame the uncertainty in global educational outcomes under alternative SSPs.

Our primary variable of interest in the article's analysis is educational attainment, which we analyze across the three SSPs. We compare forecasts resulting from the different modeling approaches of IFs and WIC (20). As previously explained, educational attainment in IFs is linked to schooling flow outcomes, with the changes coming from school completion or partial completion once attainment is initialized with historical data. So, we implement the high and low education SSP scenarios (SSP5 and SSP3, respectively) by modifying student flow rates at all levels. These flow rates are linked to demographic and economic variables as explained earlier. For a proper comparison between our schooling-based attainment projections and those from the SSP database, we also exogenously incorporate the corresponding population and economic growth projections from the SSP database into our scenarios (18).

Although IFs can respond to both student flow and schooling system financing interventions, we used only student flow parameters in this analysis. The model adjusts financing demands accordingly. As we described earlier, education spending remains subject to constraints posed by government expenditure capacity and competing public sector demands. In the SSP5 high-investment education scenario, the model ensures the largest possible education investment given the enrollment need. The specific student flow parameters used in SSP5 and SSP3 are described next.

In SSP2, student flow rates follow IFs base-case education projections, with only population and GDP growth taken from the SSP database(18). As described in Materials and Methods in the main text, for SSP5 and SSP3, two flow rates at each schooling level—one related to entry into the level (entrance/intake or transition from the level below) and one related to progression within the level (persistence to the last grade or graduation)— are adjusted to ensure that they span a wide but plausible range as compared to historical experience. The adjustment is implemented using multipliers that equal one in the base case and vary above or below one in scenarios, modifying projections of these rates subject to model constraints. Multipliers can change over time and by country; here we applied uniform values across countries, increasing gradually over time for SSP5 and decreasing gradually for SSP3. The eight student flow parameters across four schooling levels, listed in Table S2, were set to align projections with historical data as described next.

We calculated annualized changes in country-level growth rates for each entrance and progression/exit flow at all four education levels, using data from 2005–2020 but only where observations were at least 10 years apart. These growth rates were summarized in box plots to identify averages and top/bottom quartiles. The upper quartile served as the target for the median growth rate in SSP5, and the lower quartile for the median rate of decline in SSP3. Multipliers were set such that distribution (across countries) of growth rates in the student flow rates over the period 2020–2035 approximately matched the target in the historical distribution. Figure S7 shows box plots of the final growth rate distributions of student flow rates for the historical period and for the first 15 years of the three scenarios, covering all flow and enrollment rates across education levels. Table S2 reports the final multiplier values.

8. Replication

The manuscript uses International Futures System (IFs) modeling platform. IFs can be downloaded from https://ifsfiles.du.edu/. The scenarios and model run files available here contain the results used here.

The instructions for installation of IFs, running scenarios and/or loading run files and displaying results can be found in the IFs wiki(1).

References

1. Pardee Institute, International Futures (IFs) - Pardee Wiki. (No Date). Available at: https://pardeewiki.du.edu/index.php?title=International_Futures_(IFs) [Accessed 25 August 2025].

2. A. Wils, R. O'Connor, The causes and dynamics of the global education transition. (2003).

3. E. Delamonica, S. Mehrotra, J. Vandemoortele, Is EFA Affordable? Estimating the Global Minimum Cost of 'Education for All.' *Innocenti Working Paper* (2001).

4. B. Bruns, A. Mingat, R. Rakotomalala, *Achieving universal primary education by 2015: a chance for every child* (World Bank, 2003).

5. S. KC, *et al.*, Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. *Demographic Research* **22**, 383–472 (2010).

6. W. W. McMahon, *Education and Development: Measuring the Social Benefits* (Oxford University Press, 2002).

7. M. Roser, E. Ortiz-Ospina, Education Spending. *Our World in Data* (2016).

8. UNESCO Institute for Statistics (UIS), UIS Data Browser. UIS Data Browser. Deposited No Date.

9. Wittgenstein Centre, Wittgenstein Centre Human Capital Data Explorer. Deposited 2024.

10. Wittgenstein Centre, Wittgenstein Centre Human Capital Data Explorer. Deposited 2018.

11. *International Standard Classification of Education (ISCED) 2011* (UNESCO, 2012).

12. United Nations, World Population Prospects, Online Edition. Deposited 2022.

13. World Bank, *World Development Indicators* (World Bank, 2024).

14. World Bank, Education Statistics Portal.

15. International Monetary Fund, World economic outlook database.

16. R. J. Barro, J. W. Lee, A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics* **104**, 184-198 0304-3878 (2013).

17. OECD, *Education at a Glance 2023: OECD Indicators* (OECD, 2023).

18. IIASA and contributing modeling teams 2024, SSP Scenario Explorer. Deposited July 2024.

19. B. C. O'Neill, *et al.*, The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change* **42**, 169-180, (2017).

20. S. K.C., *et al.*, "Updating the Shared Socioeconomic Pathways (SSPs) Global Population and Human Capital Projections" (International Institute for Applied Systems Analysis (IIASA), 2024).

21. UNESCO Institute for Statistics, Education Non Core Archive February 2020. UIS Bulk Data Download Service. Deposited No Date.

Figures

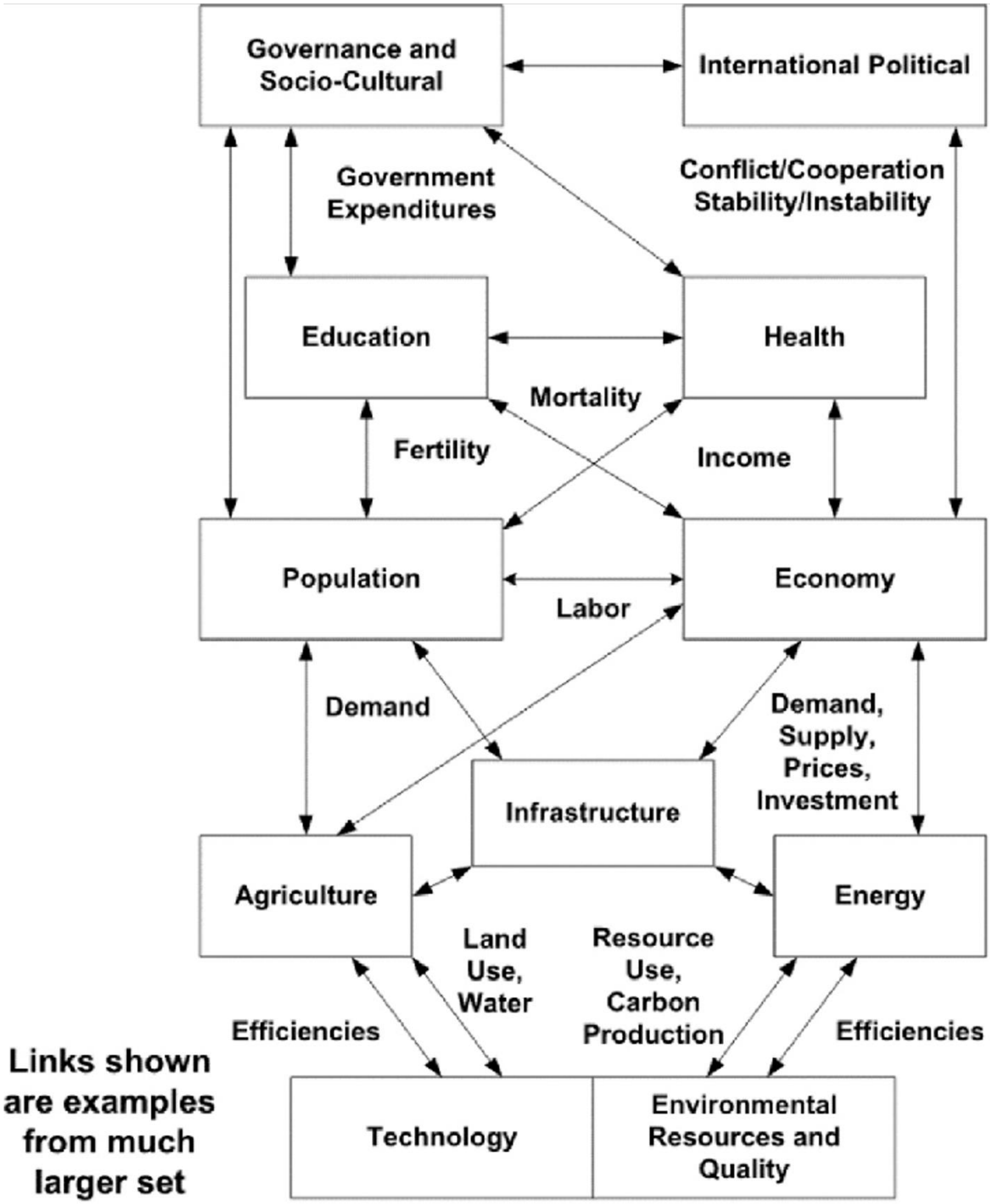


Figure S1 International Futures Models

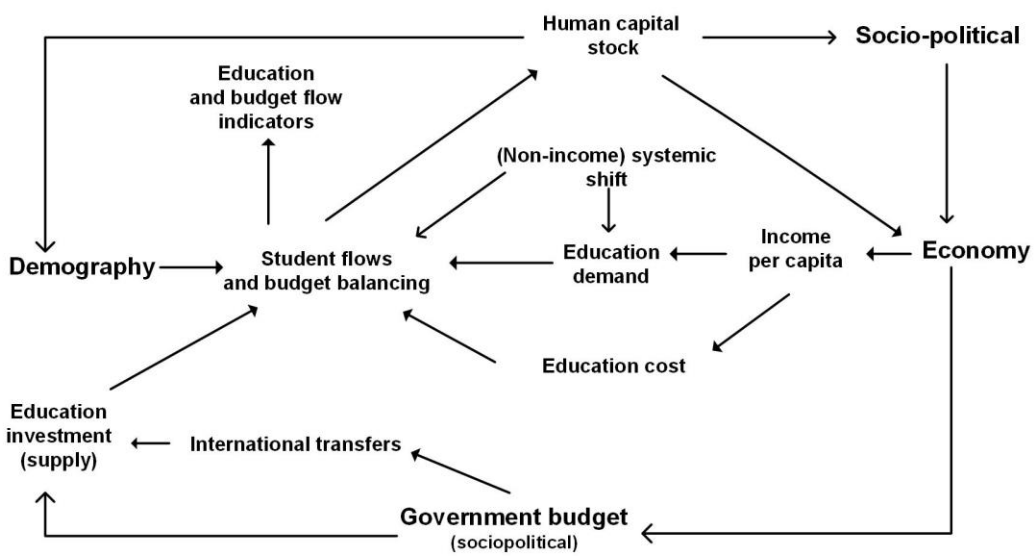


Figure S2 International Futures Education Model

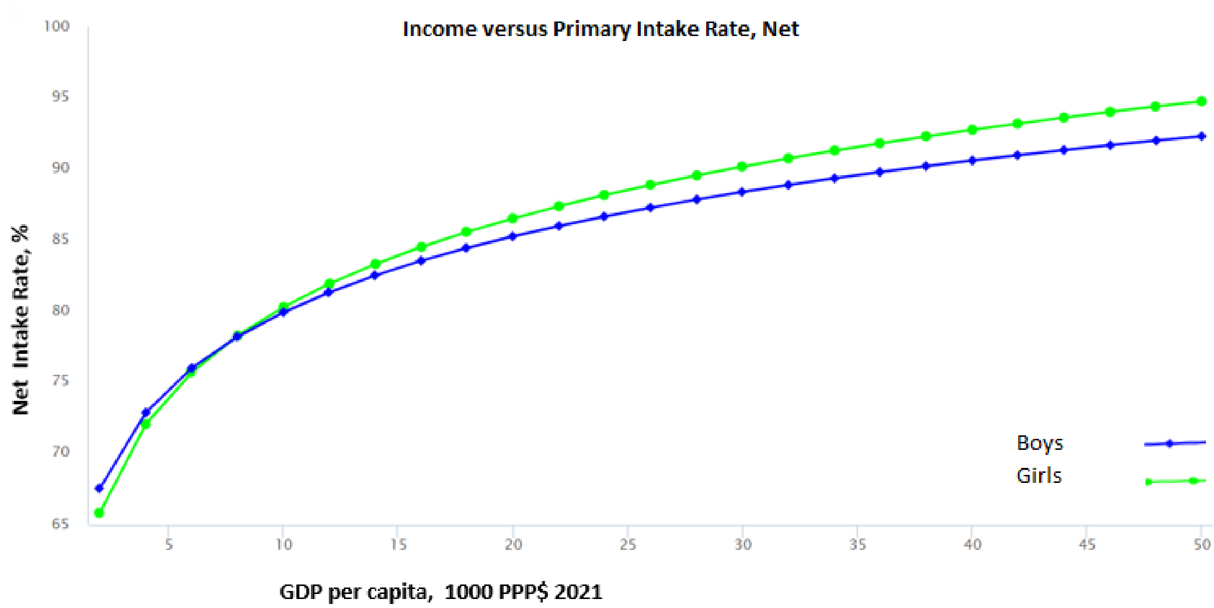


Figure S3: Plot of net intake rate in primary education (female and male) on GDP per capita (2021 PPP), using most recent data available

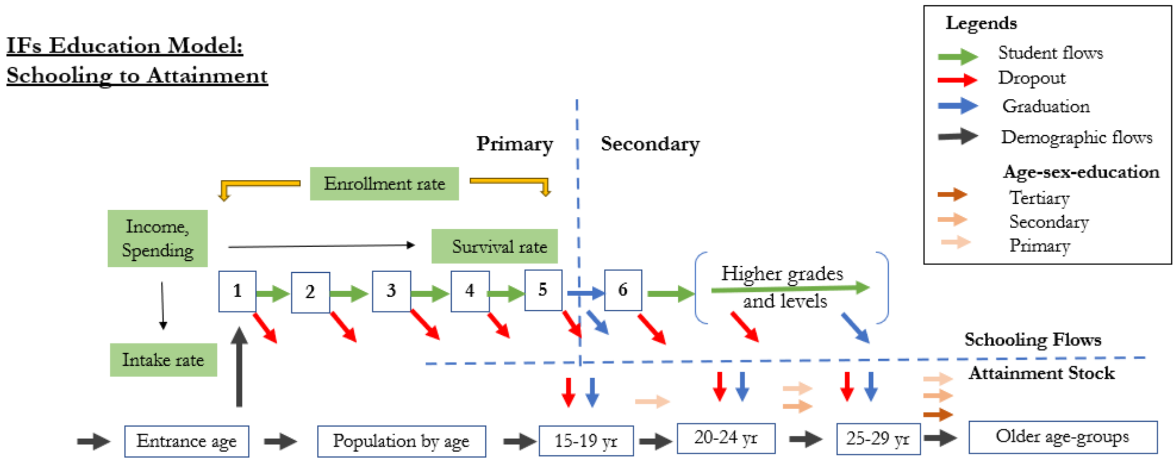


Figure S4. Schematic diagram of progress through education with linkages to drivers and adult attainment.

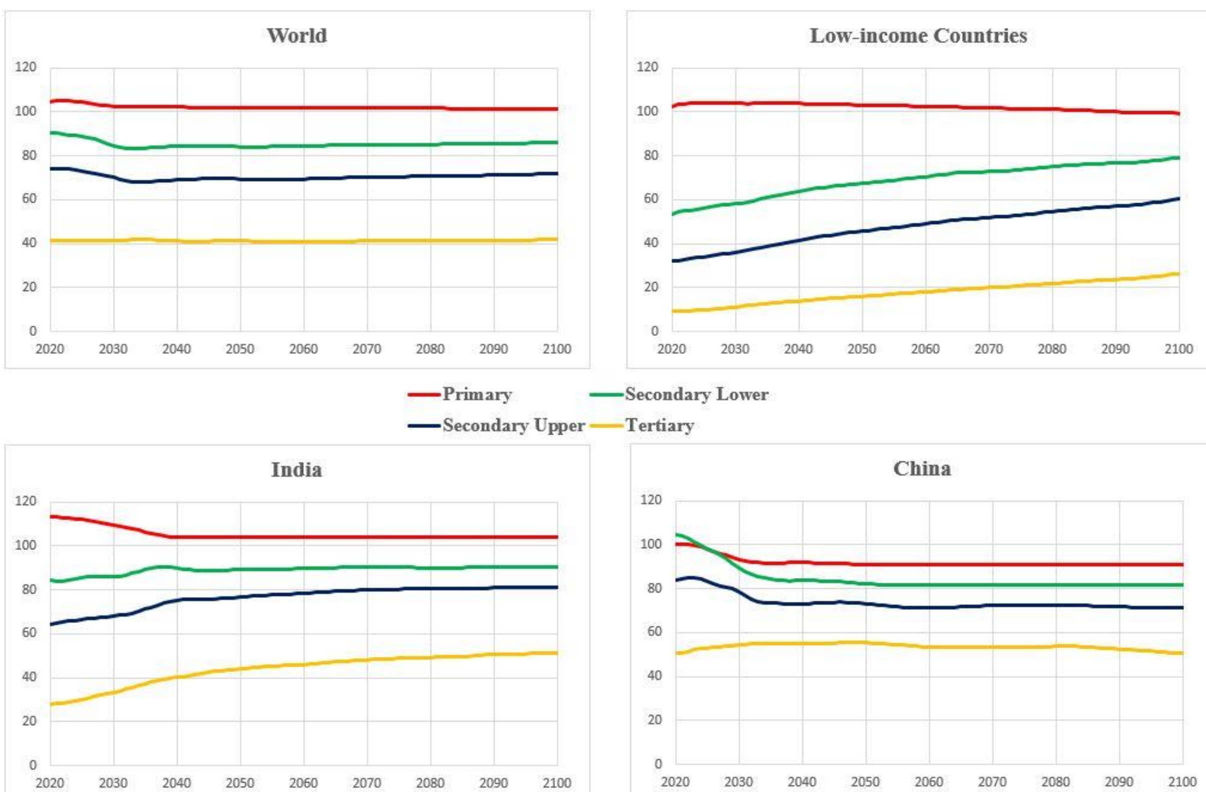


Figure S5: Gross enrollment projections for all four levels: World, LICs, China, India for SSP3

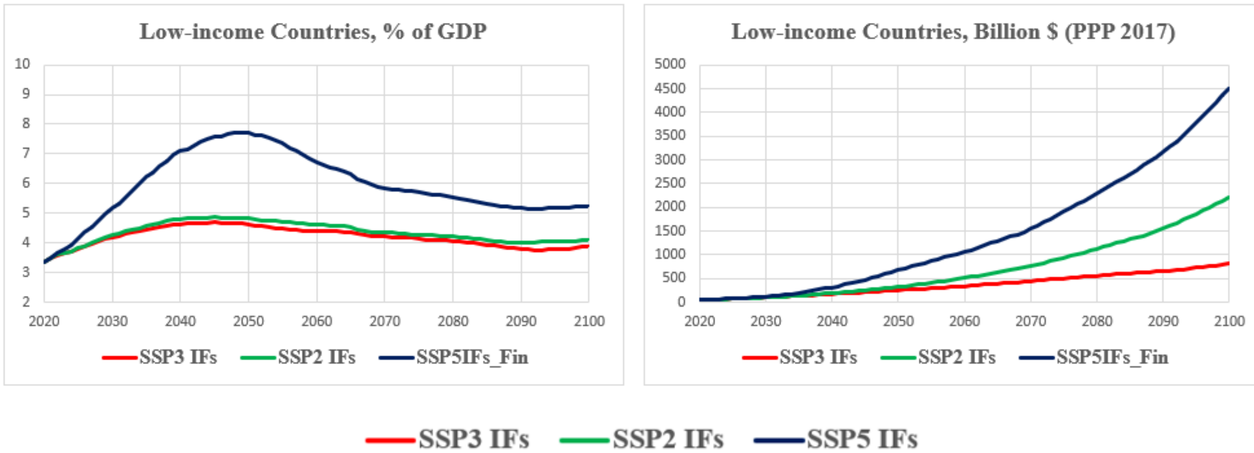
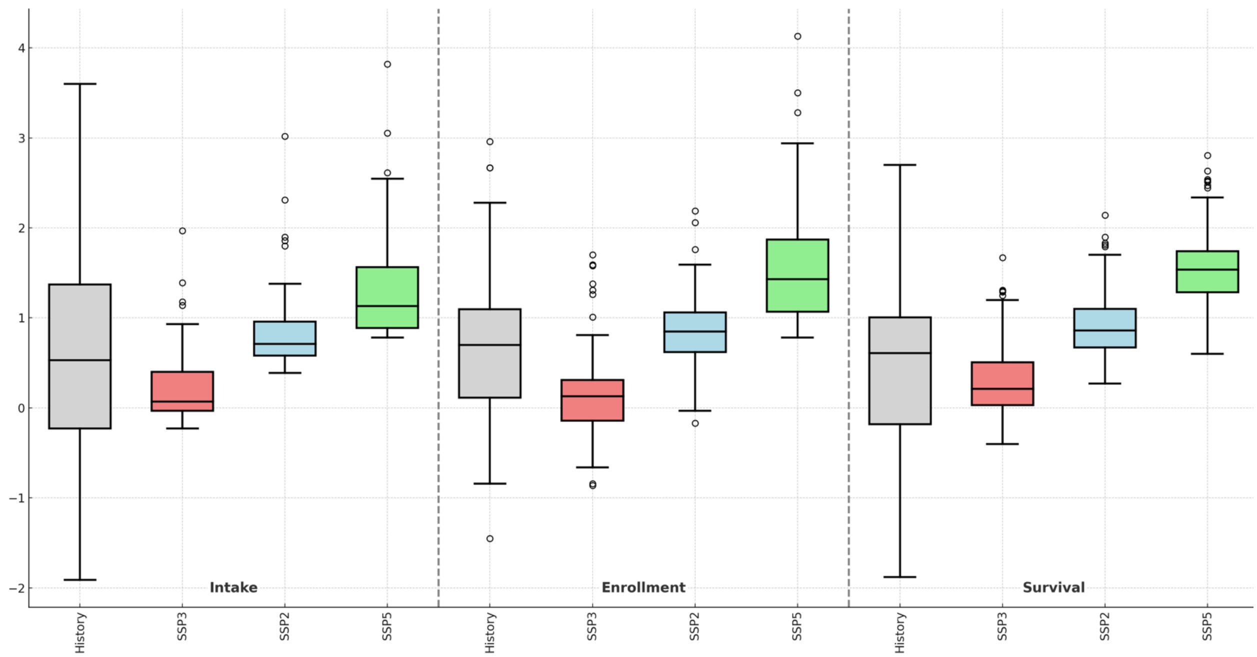
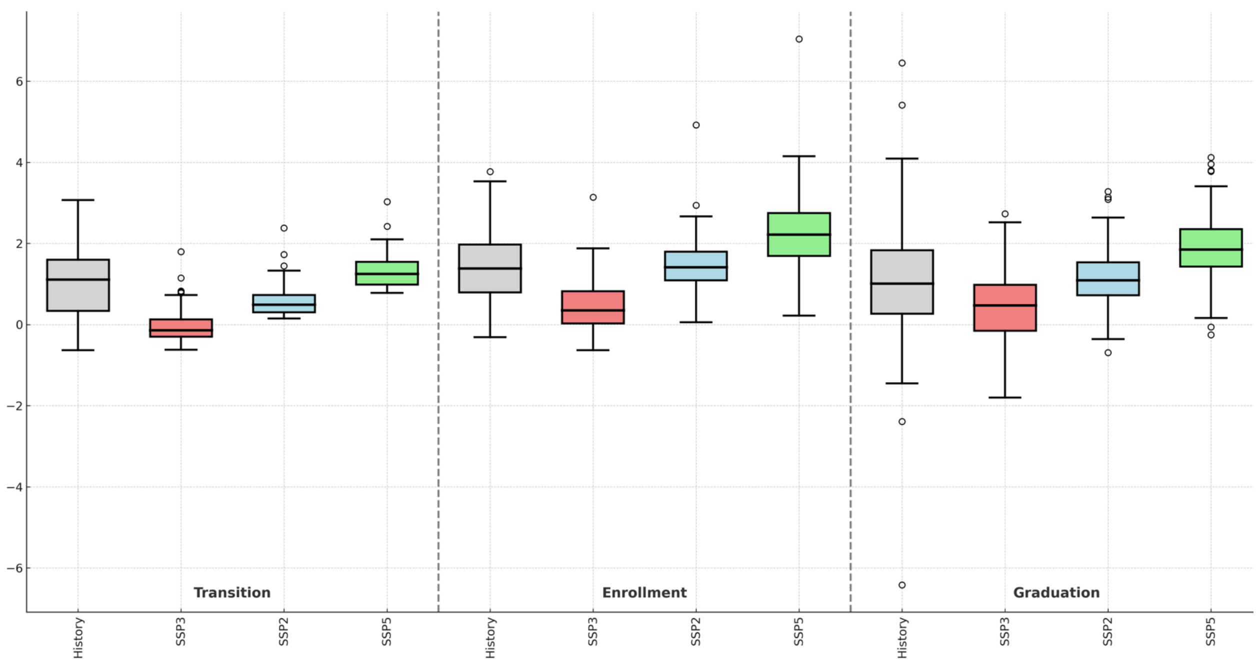


Figure S6: Government spending in education: Low-income Countries; all three SSPs

Boxplots with Growth in Student Flows, Primary Net



Boxplots with Growth in Student Flows, Secondary Lower



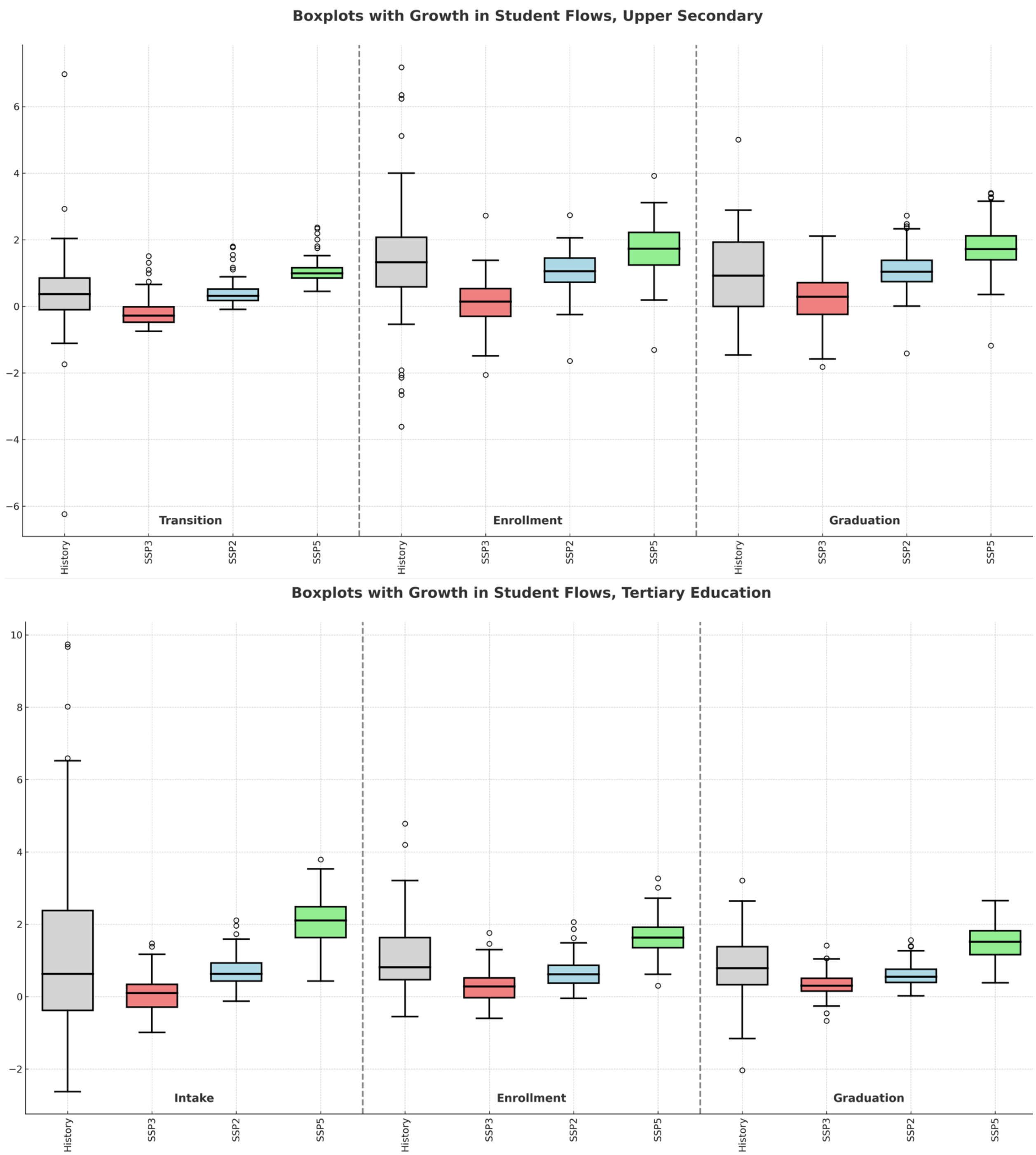


Figure S7. Boxplots of annualized growth rates in student flow rates, comparing historical data and the three SSPs

Tables

Table S1: Data sources for variables used in IFs education model

Variable	Unit	Data Source
Primary Education		
Intake Rate, primary, gross and net	% of entrance-age population	(8)
Enrollment Rate, primary, gross and net	% of relevant school-age population	(8)
Survival Rate, Primary	% of first grade entrants	(8)
Completion Rate, Primary	% of graduation-age population	(8)
Enrollment by grade, Primary	Number of students	(21)
Dropout rate by grade, Primary	% of grade enrollment	(21)
Duration of Primary	Year	(8)
Lower Secondary		
Transition Rate, Primary to Lower Secondary	% of last graders in primary	(8)
Enrollment Rate, Gross, Lower Secondary	% of relevant school-age population	(8)
Graduation Rate, Lower Secondary	% of graduation-age population	(8)
Enrollment by grade, Lower Secondary	Number of students	(21)
Dropout rate by grade, Lower Secondary	% of grade enrollment	(21)
Vocational Share, Lower Secondary	% of level enrollment	(8)
Duration of Lower Secondary	Year	(8)
Upper Secondary		
Transition Rate, Lower Sec to Upper Sec	% of last graders in lower sec	(8)
Enrollment Rate, Gross, Upper Secondary	% of relevant school-age population	(8)
Graduation Rate, Upper Secondary	% of graduation-age population	UIS (2025b)
Vocational Share, Upper Secondary	% of level enrollment	(8)
Duration of Upper Secondary	Year	(8)
Tertiary Education		
Intake Rate, Gross, Tertiary	% of entrance-age population	(8)
Enrollment Rate, Gross, Tertiary	% of relevant school-age population	(8)
Graduation Rate, Tertiary	% of graduation-age population	(8)
Education Finance		
Government Spending on Education	% of GDP	(13)
Per Student Public Expenditure (Primary, Lower Secondary, Upper Secondary, Tertiary)	% of GDP per capita	(8)
Educational Attainment		
Mean years of schooling (Age 25+, Age 15+)	Year	(9, 10)
Completion rates, adult 5-year age-groups, primary, upper secondary, tertiary	Percent of age-group	(9, 10)
Economic and Demographic		
GDP per capita at PPP	2017 PPP Dollar	(13, 15)
Population: total, by age-group	Million	(12)

Table S2: IFs SSP scenario parameters

No	Level of education	Parameter Name	SSP3 Target (Year)	SSP5 Target (Year)
1	Primary	Net intake rate	.9 (2030)	1.2 (2035)
2	Primary	Survival rate to last grade	.9 (2030)	1.2 (2035)
3	Lower Secondary	Transition rate	.9 (2030)	1.2 (2035)
4	Lower Secondary	Graduation rate	.9 (2030)	1.2 (2035)
5	Upper Secondary	Transition rate	.9 (2030)	1.2 (2035)
6	Upper Secondary	Graduation rate	.9 (2030)	1.2 (2035)
7	Tertiary	Intake rate	.8(2040)	1.4 (2035)
8	Tertiary	Graduation rate	.8(2040)	1.4 (2035)

(Note: *Parameter values are same for all countries. They are set to 1, their default value, in SSP2 (base case). In SSP3 (SSP5), they gradually decrease (increase) to the target value by the specified year and remain at that level thereafter. This reflects an effort to stagnate (accelerate) enrollment relative to the SSP2 middle-of-the-road trajectory.*