****FREDERICK S. PARDEE CENTER

FOR INTERNATIONAL FUTURES

*EXPLORE UNDERSTAND SHAPE*

WORKING PAPER 2013.06.x

INTERNATIONAL FUTURES MODEL DOCUMENTATION:

EDUCATION MODEL

Author: Mohammod Irfan

Published June 2013

Note: This is a draft version. Please do not cite or quote.



**Table of Contents**

[1. Education 3](#_Toc361928443)

[1.1 Introduction 3](#_Toc361928444)

[1.2 Structure and Agent System: Education 4](#_Toc361928445)

[1.3 Dominant Relations: Education 5](#_Toc361928446)

[1.4 Key dynamics are directly linked to the dominant relations 5](#_Toc361928447)

[1.5 Education sector: Added Value 5](#_Toc361928448)

[2 Education Model Coverage 6](#_Toc361928449)

[2.1 Education Model Coverage 6](#_Toc361928450)

[2.2 What the Model Does not Cover 7](#_Toc361928451)

[2.3 Sources of Education Data 7](#_Toc361928452)

[2.4 Reconciliation of Flow Rates 8](#_Toc361928453)

[2.5 Variable Naming Convention 8](#_Toc361928454)

[3 Education Flow Charts 9](#_Toc361928455)

[3.1 Education Overview 9](#_Toc361928456)

[3.2 Education Student Flow 10](#_Toc361928457)

[3.3 Education Attainment 12](#_Toc361928458)

[3.4 Education Financial Flows 15](#_Toc361928459)

[4 Education Equations 17](#_Toc361928460)

[4.1 Education Equations Overview 17](#_Toc361928461)

[4.2 Education Equations: Student Flow: Econometric Models for Core Inflow and Outflow 18](#_Toc361928462)

[4.3 Education Equations: Student Flow: Systemic Shift 19](#_Toc361928463)

[4.4 Education Equations: Student Flow: Grade Flow Algorithm 21](#_Toc361928464)

[4.5 Education Equations: Student Flow: Gross and Net 22](#_Toc361928465)

[4.6 Education Equations: Student Flow: Vocational Education 22](#_Toc361928466)

[4.7 Education Equations: Budget Flow 22](#_Toc361928467)

[4.8 Education Equations: Attainment 25](#_Toc361928468)

[5 Reference 28](#_Toc361928469)

# Education

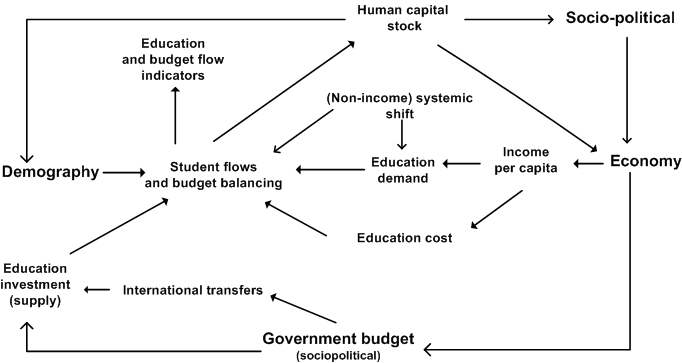
## Introduction

The education model of IFs simulates patterns of educational participation and attainment in 186 countries over a long time horizon under alternative assumptions about uncertainties and interventions (Irfan 2008). Its purpose is to serve as a generalized thinking and analysis tool for educational futures within a broader human development context.

The model forecasts gender- and country-specific access, participation and progression rates at levels of formal education starting from elementary through lower and upper secondary to tertiary. The model also forecasts costs and public spending by level of education. Dropout, completion and transition to the next level of schooling are all mapped onto corresponding age cohorts thus allowing the model to forecast educational attainment for the entire population at any point in time within the forecast horizon.

From simple accounting of the grade progressions to complex budget balancing and budget impact algorithm, the model draws upon the extant understanding and standards (e.g., UNESCO's ISCED classification explained later) about national systems of education around the world. One difference between other attempts at forecasting educational participation and attainment (e.g, McMahon 1999; Bruns, Mingat and Rakotomalala 2003; Wils and O’Connor 2003; Delamonica, Mehrotra and Vandemoortele. 2001; Cuaresma and Lutz 2007) and our forecasting, is the embedding of education within an integrated model in which demographic and economic variables interact with education, in both directions, as the model runs.

In the figure below we display the major variables and components that directly determine education demand, supply, and flows in the IFs system. We emphasize again the inter-connectedness of the components and their relationship to the broader human development system. For example, during each year of simulation, the IFs cohort-specific demographic model provides the school age population to the education model. In turn, the education model feeds its calculations of education attainment to the population model’s determination of women’s fertility. Similarly, the broader economic and socio-political systems provide funding for education, and levels of educational attainment affect economic productivity and growth, and therefore also education spending.



## Structure and Agent System: Education

|  |  |
| --- | --- |
| System/Sub System | National Education System |
| Organizing Structure | Various levels of education; Age Cohorts |
| Stocks | Educational Attainment; Enrollment |
| Flows | Intake; Graduation; Transition; Spending |
| Key Aggregate Relationships | Demand for and achievement in education changes with income, societal change  Public spending available for education rises with income level  Cost of schooling rises with income level  Lack (surplus) of public spending in education hurts (helps) educational access and progression  More education helps economic growth and reduces fertility |
| Key Agent-Class Behavior Relationships | Families send children to school; Government revenue and expenditure in education |

## Dominant Relations: Education

The dominant relationships in the model are those that determine various educational flow rates, e.g., intake rate for primary (EDPRIINT) or tertiary (EDTERINT), or survival rates in primary (EDPRISUR) or lower secondary (EDSECLOWRSUR). These rates are functions of per capita income. Non-income drivers of education are represented by upward shifts in these functions. These rates follow an S-shaped path in most cases. The flows interact with a stocks and flows structure to derive major stocks like enrollment, for the young, and attainment, for the adult.

On the financing side, the major dynamic is in the cost of education, e.g., cost per student in primary, EDEXPERPRI, the bulk of which is teachers' salary and which thus goes up with rising income.

Public spending allocation in education, GDS(Educ) is a function of national income per capita that proxies level of economic development. Demand for educational spending - determined by initial projections of enrollment and of per student cost - and total availability of public funds affect the base allocation derived from function.

For diagrams see: [Education: Student Flow](#_Education_Student_Flow); [Education Budget Flow](#_Education_Financial_Flows)

For equations see: [Student flow equation](#_Education_Equations:_Student) [Budget Equations](#_Education_Equations:_Budget)

## Key dynamics are directly linked to the dominant relations

* Intake, survival and transition rates are functions of per capita income (GDPPCP). These functions shift upward over time representing the non-income drivers of education.
* Each year flow rates are used to update major stocks like enrollment, for the young, and attainment, for the adult.
* Per student expenditure at all levels of education is a function of per capita income.
* Deficit or surplus in public spending on education, GDS(Educ) affects intake, transition and survival rates at all levels of education.

## Education sector: Added Value

IFs Education model is an integrated model. The education system in the model is interlinked with demographic, economic and socio-political systems with mutual feedback within and across theses systems. Schooling of the young is linked to education of the population as whole in this model.

The model is well suited for scenario analysis with representation of policy levers for entrance into and survival at various levels of schooling. Girls and boys are represented separately in this model.

The education budget is also endogenous o the model with income driven dynamics in cost per student for each level of education. Budget availability affect enrollment. Educational attainment raises income and affordability of education at individual and national level.

# Education Model Coverage

## Education Model Coverage

UNESCO has developed a standard classification system for national education systems called International Standard Classification of Education, ISCED. ISCED 1997 uses a numbering system to identify the sequential levels of educational systems—namely, pre-primary, primary, lower secondary, upper secondary, post-secondary non tertiary and tertiary—which are characterized by curricula of increasing difficulty and specialization as the students move up the levels. IFs education model covers primary (ISCED level 1), lower secondary (ISCED level 2), upper secondary (ISCED level 3), and tertiary education (ISCED levels 5A, 5B and 6).

The model covers 183 countries that can be grouped into any number of flexible country groupings, e.g., UNESCO regions, like any other sub-module of IFs. Country specific entrance age and school-cycle length [data are collected](#_Sources_of_Education) and used in IFs to represent national education systems as closely as possible. For all of these levels, IFs forecast variables representing student flow rates, e.g., intake, persistence, completion and graduation, and stocks, e.g., enrolment, with the girls and the boys handled separately within each country.

One important distinction among the flow rates is a gross rate versus a net rate for the same flow. Gross rates include all pupils whereas net rates include pupils who enter the school at the right age, given the statutory entrance age in the country and proceed without any repetition. The IFs education model forecasts both net and gross rates for primary education. For other levels we forecast gross rates only. It would be useful to look at the net rates at least for lower secondary, as the catch up continues up to that level. However, we could not obtain net rate data for lower secondary.

Additionally, for lower and upper secondary, the IFs model covers both general and vocational curriculum and forecasts the vocational share of total enrolment, EDSECLOWRVOC (for lower secondary) and EDSECUPPRVOC (for upper secondary). Like all other participation variables, these two are also disaggregated by gender.

The output of the national education system, i.e., school completion and partial completion of the young people, is added to the [educational attainment](#_Education_Attainment) of the adults in the population. IFs forecasts four categories of attainment - portion with no education, completed primary education, completed secondary education and completed tertiary education - separately for men and women above fifteen years of age by five year cohorts as well as an aggregate over all adult cohorts. Model software contains so-called "Education Pyramid" or a display of educational attainments mapped over five year age cohorts as is usually done for population pyramids.

Another aggregate measure of educational attainment that we forecast is the average years of education of the adults. We have several measures, EDYEARSAG15, average years of education for all adults aged 15 and above, EDYRSAG25, average years of education for those 25 and older, EDYRSAG15TO24, average years of education for the youngest of the adults aged between fifteen years to twenty four.

IFs education model also covers [financing of education](#_Education_Financial_Flows). The model forecast per student public expenditure as a share of per capita income. The model also forecast total public spending in education and the share of that spending that goes to each level of education.

## What the Model Does not Cover

ISCED level 0, pre-primary, and level 4, post-secondary pre tertiary, are not common across all countries and are thus excluded from IFs education model.

On the financing side, the model does not include private spending in education, a significant share of spending especially for tertiary education in many countries and even for secondary education in some countries. Scarcity of good data and lack of any pattern in the historical unfolding precludes modelling private spending in education.

Quality of national education system can also vary across countries and over time. The IFs education model does not forecast any explicit indicator of education quality. However, the survival and graduation rates that the model forecasts for all levels of education are implicit indicators of system quality. At this point IFs does not forecast any indicator of cognitive quality of learners. However, the IFs database does have data on cognitive quality.

The IFs education model does not cover private spending in education.

## Sources of Education Data

UNESCO is the UN agency charged with collecting and maintaining education-related data from across the world. UNICEF collects some education data through their MICS survey. USAID also collects education data as a part of its Demographic and Household Surveys (DHS). OECD collects better data especially on tertiary education for its members as well as few other countries.

We collected our [student flows](#_Education_Student_Flow) and per student cost data from UNESCO Institute for Statistics' (UIS) web data repository at <http://stats.uis.unesco.org/unesco/tableviewer/document.aspx?ReportId=143>. (Accessed on 05/17/2013)

For [educational attainment](#_Education_Attainment) data we use estimates by Robert Barro and Jong Wha Lee (2000). They have published their estimates of human capital stock (i.e., the educational attainment of adults) at the website of the Center for International Development of Harvard University. In 2001, Daniel Cohen and Marcelo Soto presented a paper providing another human capital dataset for a total of ninety-five countries. We collect that data as well in our database.

When needed we also calculated our own series using underlying data from UNESCO. For example, we calculate an adjusted net intake rate for primary using the age specific intake rates that UNESCO report. We also calculated survival rates in lower and upper secondary (EDSECLOWRSUR, EDSECUPPRSUR) using a reconstructed cohort simulation method from grade-wise enrollment data for two consecutive years. The transition rate from lower to upper secondary is also calculated using grade data.

## Reconciliation of Flow Rates

Incongruities among the base year primary flow rates (intake, survival, and enrollment) can arise either from reported data values that, in combination, do not make sense, or from the use of “stand-alone” cross-sectional estimations used in the IFs pre-Processor to fill missing data. Such incongruities might arise among flow rates within a single level of education (e.g., primary intake, survival, and enrollment rates that are incompatible) or between flow rates across two levels of education (e.g., primary completion rate and lower secondary intake rate).

The IFs education model uses algorithms to reconcile incongruent flow values. They work by (1) analyzing incongruities; (2) applying protocols that identify and retain the data or estimations that are probably of higher quality; and (3) substituting recomputed values for the data or estimations that are probably of lesser quality. For example, at the primary level, data on enrollment rates are more extensive and more straight-forward than either intake or survival data; in turn, intake rates have fewer missing values and are arguably more reliable measures than survival rates. The IFs pre-processor reconciles student flow data for Primary by using an algorithm that assumes enrollment numbers to be more reliable than the entrance data and entrance data to be more reliable than survival data.

## Variable Naming Convention

All education model variable names start with a two-letter prefix of 'ED' followed, in most cases, by the three letter level indicator - PRI for primary, SEC for secondary, TER for tertiary. Secondary is further subdivided into SECLOWR for lower secondary and SECUPPR for upper secondary. Parameters in the model, which are named using lowercase letters like those in other IFs modules, also follow a similar naming convention.

# Education Flow Charts

## Education Overview

For each country, the IFs education model represents a multilevel formal education system that starts at primary and ends at tertiary. [Student flows](#_Education_Student_Flow), i.e., entry into and progression through the system are determined by forecasts on intake and persistence (or survival) rates superimposed on the population of the corresponding age cohorts obtained from IFs population forecasts. Students at all levels are disaggregated by gender. Secondary education is further divided into lower and upper secondary, and then further into general and vocational according to the curricula that are followed.

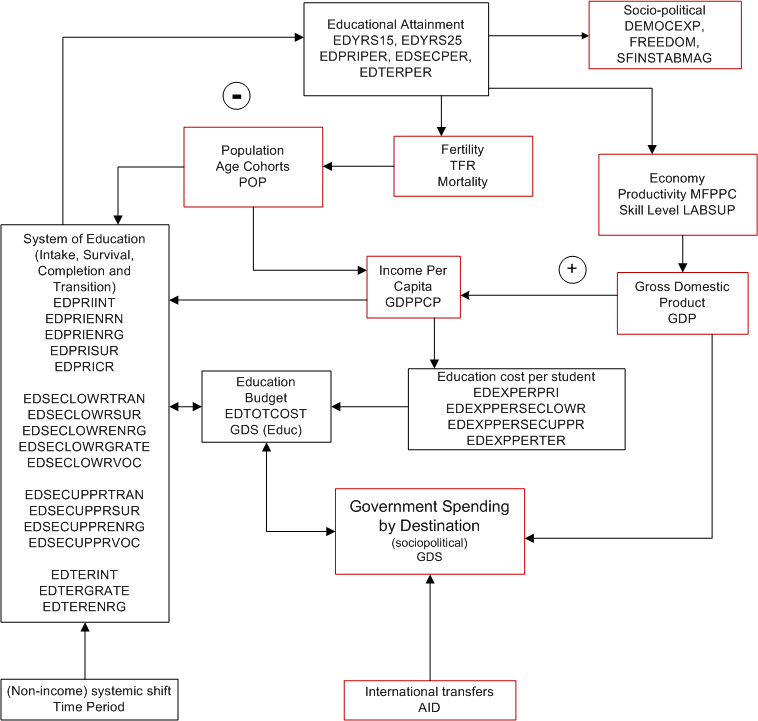
The model represents the dynamics in [education financing](#_Education_Financial_Flows) through per student costs for each level of education and a total public spending in education. Policy levers are available for changing both spending and cost.

School completion (or dropout) in the education model is carried forward as the [educational attainment](#_Education_Attainment) of the overall population. As a result, the education model forecasts population structures by age, sex, and attained education, i.e., years and levels of completed education.

The major agents represented in the education system of the model are households,—represented by the parents who decide which of their boys and girls will go to school—and governments that direct resources into and across the educational system. The major flows within the model are student and budgetary, while the major stock is that of educational attainment embedded in a population. Other than the budgetary variables, all the flows and stocks are gender disaggregated.

The education model has forward and backward linkages with other parts of the IFs model. During each year of simulation, the IFs cohort-specific demographic model provides the school age population to the education model. In turn, the education model feeds its calculations of education attainment to the population model’s determination of women’s fertility. Similarly, the broader economic and socio-political systems provide funding for education, and levels of educational attainment affect economic productivity and growth, and therefore also education spending.

The figure below shows the major variables and components that directly determine education demand, supply, and flows in the IFs system. The diagram attempts to emphasize on the inter-connectedness of the education model components and their relationship to the broader human development system.

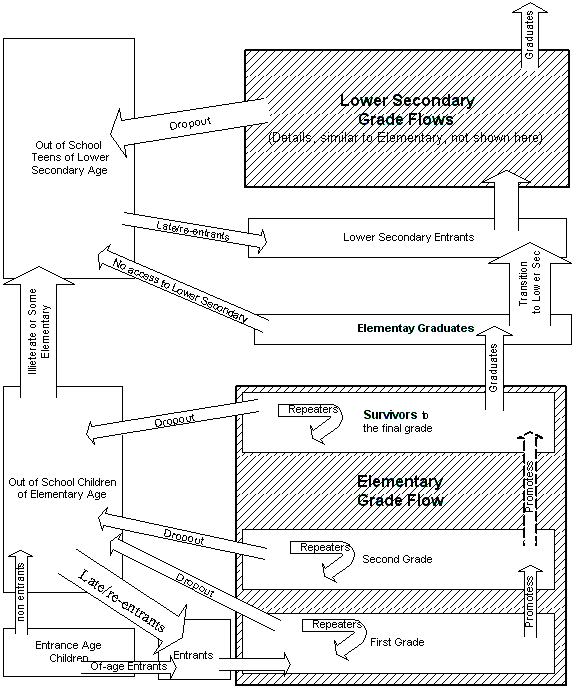


## Education Student Flow

IFs education model simulates grade-by-grade student flow for each level of education that the model covers. Grade-by-grade student flow model combine the effects of grade-specific dropout, repetition and reentry into an average cohort-specific *grade-to-grade flow rate*, calculated from the survival rate for the cohort. Each year the number of new entrants is determined by the forecasts of the intake rate and the entrance age population. In successive years, these entrants are moved to the next higher grades, one grade each year, using the *grade-to-grade flow rate*. The simulated grade-wise enrollments are then used to determine the total enrollment at the particular level of education. Student flow at a particular level of education, e.g., primary, is culminated with rates of completion and transition by some to the next level, e.g., lower secondary.

The figure below shows details of the student flow for primary (or, elementary) level. This is illustrative of the student flow at other levels of education. We model both net and gross enrollment rates for primary. The model tracks the pool of potential students who are above the entrance age (as a result of never enrolling or of having dropped out), and brings back some of those students, marked as late/reentrant in the figure, (dependent on initial conditions with respect to gross versus net intake) for the dynamic calculation of total gross enrollments.

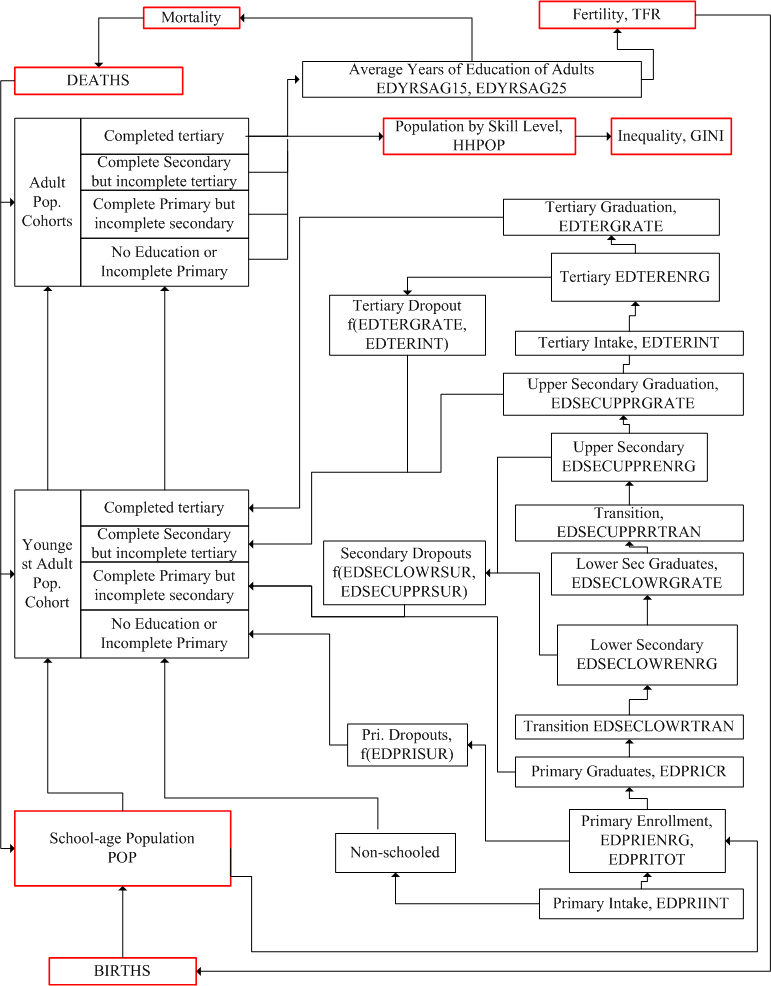
A generally similar grade-flow methodology models lower and upper secondary level student flows. We use country-specific entrance ages and durations at each level. As the historical data available does not allow estimating a rate of transition from upper secondary to tertiary, the tertiary education model calculates a tertiary intake rate from tertiary enrollment and graduation rate data using an algorithm which derives a tertiary intake with a lower bound slightly below the upper secondary graduation rate in the previous year.



## Education Attainment

The algorithm for the tracking of education attainment is very straight-forward. The model maintains the structure of the population not only by age and sex categories, but also by years and levels of completed education. In each year of the model’s run, the youngest adults pick up the appropriate total years of education and specific levels of completed education. The model advances each cohort in 1-year time steps after subtracting deaths. In addition to cohort attainment, the model also calculates overall attainment of adults (15+ and 25+) as average years of education (EDYRSAG15, EDYRSAG25) and as share of people 15+ with a certain level of education completed (EDPRIPER, EDSECPER, EDTERPER).

One limitation of our model is that it does not represent differential mortality rates associated with different levels of education attainment (generally lower for the more educated). [[1]](#footnote-1) This leads, other things equal, to a modest underestimate of adult education attainment, growing with the length of the forecast horizon. The averaging method that IFs uses to advance adults through the age/sex/education categories also slightly misrepresents the level of education attainment in each 5-year category.



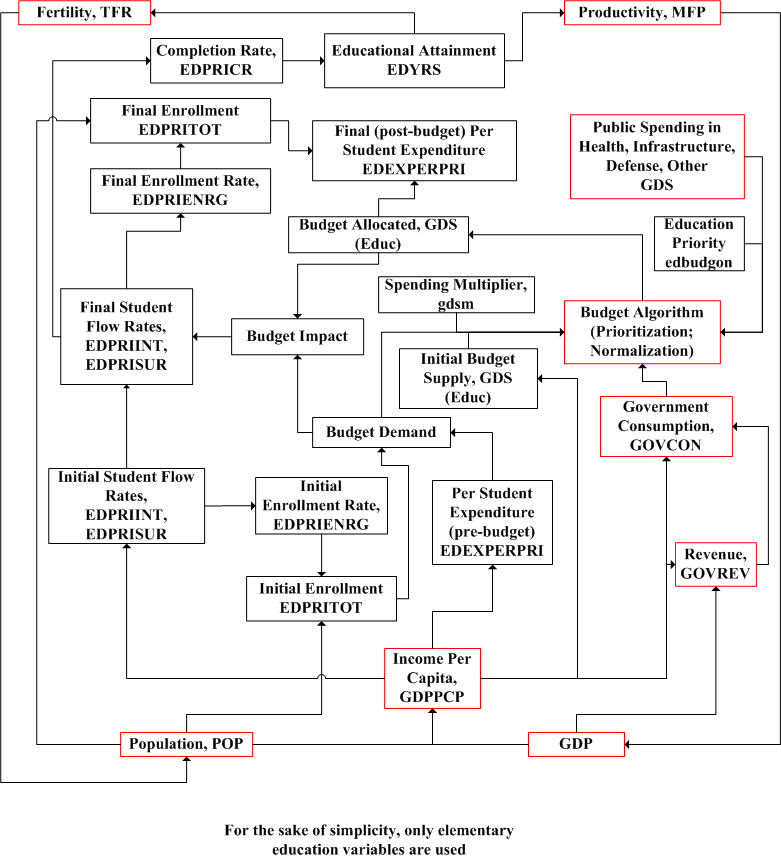
## Education Financial Flows

In addition to [student flows](#_Education_Student_Flow), and interacting closely with them, the IFs education model also tracks financing of education. Because of the scarcity of private funding data, IFs specifically represents public funding only, and our formulations of public funding implicitly assume that the public/private funding mix will not change over time.

The accounting of educational finance is composed of two major components, per student cost and the total number of projected students, the latter of the two is discussed in the [student flows](#_Education_Student_Flow) section. Spending per student at all levels of education is driven by average income. Given forecasts of spending per student by level of education and given initial enrollments forecasts by level, an estimate of the total education funding demanded is obtained by summing across education levels the products of spending per student and student numbers.

The funding needs are sent to the IFs sociopolitical model where educational spending is initially determined from the patterns in such spending regressed against the level of economic development of the countries. A priority parameter (**edbudgon**) is then used to prioritize spending needs over spending patterns. This parameter can be changed by model user within a range of values going from zero to one with the zero value awarding maximum priority to fund demands. Finally, total government consumption spending (GOVCON) is distributed among education and other social spending sectors, namely infrastructure, health, public R&D, defense and an "other" category, using a normalization algorithm.

Government spending is then taken back to the education module and compared against fund needs. Budget impact, calculated as a ratio of the demanded and allocated funds, makes an impact on the initial projection of student flow rates (intake, survival, and transition). The positive (upward) side of the budget impact is non-linear with the maximum boost to growth occurring when a flow rate is at or near its mid-point or within the range of the inflection points of an assumed S-shaped path, to be precise. Impact of deficit is more or less linear except at impact ratios close to 1, whence the downward impact is dampened. Final student flow rates are used to calculate final enrollment numbers using population forecasts for relevant age cohorts. Finally, cost per students are adjusted to reflect final enrollments and fund availability.



# Education Equations

## Education Equations Overview

The IFs education model represent two types of educational stocks, [stocks of pupils](#_Education_Equations:_Student) and stocks of adults with a certain level of [educational attainment](#_Education_Equations:_Attainment). These stocks are initialized with historical data. The simulation model then recalculates the stock each year from its level the previous year and the net annual change resulting from inflows and outflows.

The core dynamics of the model is in these [flow rates](#_Education_Equations:_Student). These flow rates are expressed as a percentage of age-appropriate population and thus have a theoretical range of zero to one hundred percent. Growing systems with a saturation point usually follow a sigmoid (S-shaped) trajectory with low growth rates at the two ends as the system begins to expand and as it approaches saturation. Maximum growth in such a system occurs at an inflection point, usually at the middle of the range or slightly above it, at which growth rate reverses direction. Some researchers (Clemens 2004; Wils and O’Connor 2003) have identified sigmoid trends in educational expansion by analyzing enrollment rates at elementary and secondary level. The IFs education model is not exactly a trend extrapolation; it is rather a forecast based on fundamental drivers, for example, income level. Educational rates in our model are driven by income level, a systemic shift algorithm and a [budget impact](#_Education_Equations:_Budget) resulting from the availability of public fund. However, there are growth rate parameters for most of the flows that allow model user to simulate desired growth that follows a sigmoid-trajectory. Another area that makes use of a sigmoid growth rate algorithm is the boost in flow rates as a result of budget surplus.

Intake (or transition), survival, enrollment and completion are some of the rates that IFs model forecast. Rate forecasts [cover](#_Overview_of_Education) elementary, lower secondary, upper secondary and tertiary levels of education with separate equations for boys and girls for each of the rate variables. All of these rates are required to calculate pupil stocks while completion rate and dropout rate (reciprocal of survival rate) are used to determine educational attainment of adults.

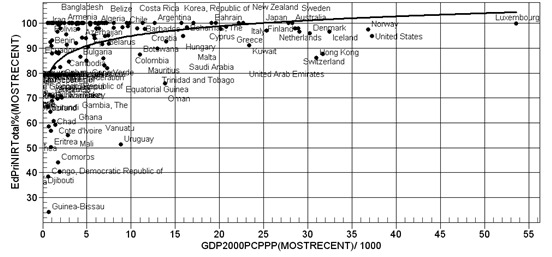
On the financial side of education, IFs forecast cost per student for each level. These per student costs are multiplied with enrollments to calculate fund demand. Budget allocation calculated in IFs socio-political module is sent back to education model to calculate final enrollments and cost per student as a result of fund shortage or surplus.

The population module provides cohort population to the education model. The economic model provides per capita income and the socio-political model provides budget allocation. Educational attainment of adults calculated by the education module affects fertility and mortality in the population and health modules, affects productivity in the economic module and affects other socio-political outcomes like governance and democracy levels.

## Education Equations: Student Flow: Econometric Models for Core Inflow and Outflow

Enrollments at various levels of education - EDPRIENRN, EPRIENRG, EDSECLOWENRG, EDSECUPPRENRG, EDTERENRG - are initialized with historical data for the beginning year of the model. Net change in enrollment at each time step is [determined by inflows](#_Education_Student_Flow) (intake or transition) and outflows (dropout or completion). Entrance to the school system (EDPRIINT, EDTERINT), transition from the lower level (EDSECLOWRTRAN, EDSECUPPRTRAN) - and outflows - completion (EDPRICR), dropout or it's reciprocal, survival (EDPRISUR) - are some of these rates that are forecast by the model.

The educational flow rates are best explained by per capita income that serves as a proxy for the families' opportunity cost of sending children to school. For each of these rates, separate regression equations for boys and girls are estimated from historical data for the most recent year. These regression equations, which are updated with most recent data as the model is rebased with new data every five years, are usually logarithmic in form. The following figure shows such a regression plot for net intake rate in elementary against per capita income in PPP dollars.



In each of the forecast years, values of the educational flow rates are first determined from these regression equation. Independent variables used in the regression equations are endogenous to the IFS model. For example, per capita income, GDPPCP, forecast by the IFs economic model drives many of the educational flow rates. The following equation shows the calculation of one such student flow rate (CalEdPriInt) from the log model of net primary intake rate shown in the earlier figure.

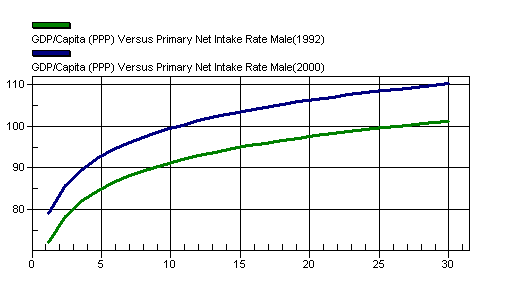
Subscript g in this equation (and all other equations of this module) stands for gender (with a value of 1 for boys, 2 for girls and 3 for boys and girls combined), r stands for countries and t stands for time period in year.

While all countries are expected to follow the regression curve in the long run, the residuals in the base year make it difficult to generate a smooth path with a continuous transition from historical data to regression estimation. We handle this by adjusting regression forecast for country differences using an algorithm that we call "shift factor" algorithm. In the first year of the model run we calculate a shift factor (EDPriIntNShift)as the difference (or ratio) between historical data on net primary intake rate (EDPRIINTN) and regression prediction for the first year for all countries. As the model runs in subsequent years, these shift factors (or initial ratios) converge to zero or one if it is a ratio (code routine ConvergeOverTime in the equation below) making the country forecast merge with the global function gradually. The period of convergence for the shift factor (PriIntN\_Shift\_Time) is determined through trial and error in each case.

The base forecast on flow rates resulting from of this regression model with country shift is used to calculate the demand for funds. These base flow rates might change as a result of budget impact based on the availability or shortage of education budget explained in the [budget flow section](#_Education_Equations:_Budget).

## Education Equations: Student Flow: Systemic Shift

Access and participation in education increases with socio-economic developments that bring changes to people's perception about the value of education. This upward shifts are clearly visible in cross-sectional regression done over two adequately apart points in time. The next figure illustrates such shift by plotting net intake rate for boys at the elementary level against GDP per capita (PPP dollars) for two points in time, 1992 and 2000.



IFs education model introduces an algorithm to represent this shift in the regression functions. This "systemic shift" algorithm starts with two regression functions about 10 to 15 years apart. An additive factor to the flow rate is estimated each year by calculating the flow rate (CalEdPriInt1 and CalEdPriInt2 in the equations below) progress required to shift from one function, e.g., to the other, s, in a certain number of years (SS\_Denom), as shown below. This systemic shift factor (CalEdPriIntFac) is then added to the flow rate (EDPRIINTN in this case) for a particular year (t) calculated from regression and country shift as described in the previous section.

As said earlier, [Student flow](#_Education_Student_Flow) rates are expressed as a percentage of underlying stocks like the number of school age children or number of pupils at a certain grade level. The flow-rate dynamics work in conjunction with population dynamics (modeled inside IFs population module) to forecast enrollment totals.

## Education Equations: Student Flow: Grade Flow Algorithm

Once the core inflow (intake or transition) and outflow (survival or completion) are determined, enrollment is calculated from grade-flows. Our grade-by-grade student flow model therefore uses some simplifying assumptions in its calculations and forecasts. We combine the effects of grade-specific dropout, repetition and reentry into an average cohort-specific grade-to-grade dropout rate, calculated from the survival rate (EDPRISUR for primary) of the entering cohort over the entire duration of the level (**EDPRILEN** for primary). Each year the number of new entrants is determined by the forecasts of the intake rate (EDPRIINT) and the entrance age population. In successive years, these entrants are moved to the next higher grades, one grade each year, subtracting the grade-to-grade dropout rate (DropoutRate). The simulated grade-wise enrollments (GradeStudents with Gcount as a subscript for grade level) are then used to determine the total enrollment at the particular level of education (EDPRIENRG for Primary).

There are some obvious limitations of this simplified approach. While our model effectively includes repeaters, we represent them implicitly (by including them in our grade progression) rather than representing them explicitly as a separate category. Moreover, by setting first grade enrollments to school entrants, we exclude repeating students from the first grade total. On the other hand, the assumption of the same grade-to-grade flow rate across all grades might somewhat over-state enrollment in a typical low-education country, where first grade drop-out rates are typically higher than the drop-out rates in subsequent grades. Since our objective is to forecast enrollment, attainment and associated costs by level rather than by grade, however, we do not lose much information by accounting for the approximate number of school places occupied by the cohorts as they proceed and focusing on accurate representation of total enrollment.

(Note: EDPRILEN is in lowercase in IFsVar)

## Education Equations: Student Flow: Gross and Net

Countries with a low rate of schooling, especially those that are catching up, usually have a large number of over-age students. Enrollment and entrance rates that count students of all ages are called gross rates in contrast to the net rate that only takes the of-age students in the numerator of the rate calculation expression. UNESCO report net and gross rates separately for entrance and participation in elementary. IFs education model forecasts both net and gross rate in primary education. An overage pool (PoolPrimary) is estimated at the model base year using net and gross intake rate data. Of-age non-entrants continue to add to the pool (PoolInflow). The pool is exhausted using a rate (PcntBack) determined by the gross and net intake rate differential at the base year. The over-age entrants (cOverAgeIntk\_Pri) gleaned from the pool are added to the net intake rate (EDPRIINTN) to calculate the gross intake rate (EDPRIINT).

## Education Equations: Student Flow: Vocational Education

For lower and upper secondary, IFs education model forecasts vocational education (EDSECLOWRVOC, EDSECUPPRVOC). Participation in vocational education depends on the policy decisions made by the country and is thus difficult to forecast. We use a simple algorithm of maintaining the vocational share of total enrollment unless changed through policy intervention. At this point, we do not have any parameter to change vocational shares. We plan to add such parameters in the future.

## Education Equations: Budget Flow

Resources required to maintain the projected student flows are determined by multiplying enrollment rates with per student cost forecasts. Availability of resources, as determined in the IFs socio-political model, affect flow rates and the final enrollment rate.

Public expenditure per student (EDEXPERPRI) as a percentage of per capita income is first estimated (CalExpPerStud) using a regression equation. Country situations are added as a shift factor (EdExPerPriShift) that wears off over a period of time (**edexppconv**) in the same manner as those for student flow rates. The following group of equations show the calculation of per student expenditure in primary (EDEXPERPRI).

Total fund demand (EDBUDDEM, see calculation below) is passed to the IFs socio-political model where a detail government budget model distributes total government consumption among various public expenditure sectors. For education allocation, an initial estimate (gkcomp) is first made from a regression function of educational spending as a percentage of GDP over GDP per capita at PPP dollars (GDPPCP) as a country gets richer.

Like several other functions discussed in this sub-module, country situation is reflected by estimating country ratio (gkri) between the predicted and historical value in the base year. This ratio converges to a value of one very slowly essentially maintaining the historic ratio. Public spending on education in billion dollars (GDS) is then calculated using the regression result, GDP and the multiplicative shift.

Sociopolitical model also forecast public spending in other areas of social spending, i.e., military, health, R&D. Another public spending sector, infrastructure is calculated bottom-up, i.e., as an aggregation of demand for construction and maintenance of various types of infrastructure.

Once all the spending shares are projected, a normalization algorithm is used to distribute the total available government consumption budget (GOVCON) among all sectors.

Before normalization, a priority parameter allows setting aside all or part of fund demands for the ground up spending sectors, i.e., infrastructure and education. For education sector, the prioritization parameter (**edbudgon**) is used to set aside a certain portion of the projected education investment as shown in the equations below.

Education allocation, GDS (Educ) calculated thus is taken back to the education model. A second normalization and prioritization is done within the education model to distribute total education allocation among different levels of education. This across level normalization uses the percentage share of each educational level in the total demand for education funding. First, total expenditure demand for all levels of education combined is determined by multiplying the total enrollments with per student costs. The following equation shows the calculation for Primary (Note: Will have to take the underscores off from code).

Fund demands for all levels are added up to get the total fund demand under no budget constraint. The prefixes UD here stands for budget unconstrained demand

Any surplus or deficit in educational allocation, calculated as the difference between education sector allocation in the government budget model and the total fund requirement for all levels of education combined, first undergoes an adjustment algorithm that boosts (in case of surplus) or reduces (in case of deficit) per student cost for those countries which are below or above the level they are supposed to be. Post this adjustment, allocation is distributed across all levels using a normalization process based on demand.

A budget impact ratio is then calculated as the ratio of the fund demanded (CalcTotCost) and fund obtained (CalcTotSpend). This budget impact ratio (CalcBudgetImpact) increases or decreases the pre-budget (or demand side as we call it) projection of [student flow rates](#_Education_Equations:_Student)  (intake, survival, and transition). The positive (upward) side of the budget impact is non-linear with the maximum boost to growth occurring when a flow rate is at or near its mid-point or within the range of the inflection points of an assumed S-shaped path, to be precise. Impact of deficit is more or less linear except at impact ratios close to 1, whence the downward impact is dampened. Final student flow rates are used to calculate final enrollment numbers using population forecasts for relevant age cohorts. Finally, cost per students are adjusted to reflect final enrollments and fund availability.

Budget impacts uses a non-linear algorithm intended to generate an S-shaped growth rate. Final enrollment is then calculated from this final flow rates and any of the remaining budget is used to increase per student expenditure.

In the equations above, convtoexchange is a factor that converts monetary units from PPP to exchange rate dollars, SpendCostRI is a ratio calculated at the first year of the model to reconcile historical data on aggregate and bottom-up spending.

## Education Equations: Attainment

There are two types of variables that keep track of educational attainment: average years of education of adults (EDYRSAG15, EDYRSAG15TO24 and EDYRSAG25) and percentage of adults with a certain level of education (EDPRIPER, EDSECPER, EDTERPER). Both groups forecast attainment by gender.

The basis of calculation for both groups of variables is educational attainment by age cohort and gender as contained in intermediate model variables, EDPriPopPer r.g,c,t , EDSecPopPer r.g,c,t, EdTerPopPer r.g,c,t (where, r stands for country or region, g for gender, c for cohort and t for time).

We initialize attainments of the entire adult population (EDPRIPER, EDSECPER, EDTERPER) using historical data estimated by Barro and Lee (2000) and use a spread algorithm. The spread algorithm starts with the most recent data on school completion rate (EDPRICR for primary) which is considered as the average attainment of the graduating cohort. The algorithm then uses the differential between that completion rate and the attainment rate of the adults (EDPRIPER) to back calculates a delta reduction for each of the older cohorts (EdPriPopPer) such that averaging attainments over cohorts one can obtain average attainment for all adults (EDPRIPER

where, subscript c stand for five year age cohorts going from 1 to 21. Cohort 4, represents the 15 to 19 years and NC, total number of age cohorts.

For subsequent forecast years, cohort educational attainment for each level of education is calculated by adding graduates from that level of education to the appropriate age cohort, advancing graduates from the younger cohort, and passing graduates to the older cohort.

where, pc stands for the five year age cohort where the primary graduates belong. For all other cohorts:

Cohort attainments for secondary and tertiary education (EDSECPOPPER, EDTERPOPPER) are initialized and forecast in a similar fashion. An average years of education reflecting completion of levels is then calculated by from the cohort attainment, population and cohort length as shown in the next equation where AGEDSTc,g,r,t contains the population of five year age cohorts and **EDPRILEN**r,t is the duration of primary cycle in years.

(Note: EDPRILEN is in lowercase in IFsVar; AGEDST is in all caps in other places)

For those who dropout before completing a certain level we need to calculate the partial attainment and add that to the average years of education. The average of the partial years of education at a particular year is calculated from dropouts by level and grade as shown below. Calculation of the average of partial years resulting from dropouts in primary education is illustrated in the equations below. Partial years from current year dropouts at other levels of education are calculated in the same manner and all the partial years are averaged to an overall average. This new partial attainment is then added to the partial attainment of five year cohorts which are initialized and advanced in a similar manner as that used for cohort averages on completed attainment.

Here, EDPRISUR is the survival rate in primary education, EDPRISTART is the official entrance age for primary schooling, Gr\_Students is the enrollment at a certain grade, GCount is the grade counter and FAGEDST is the population of the single year age cohort corresponding to the grade level.

(Note: EDPRILEN and EDPRISTART are in lowercase in IFsVar; need to change)

Overall attainment, i.e., average years of education are calculated by averaging the attainments and partial attainments of five year age cohorts as shown in the equation below. The suffixes on the variables EDYRSAG15, EDYRSAG15TO24 and EDYRSAG25 indicate the age thresholds at which or the age bracket over which attainment is averaged.

Attainments by level, i.e., EDPRIPER, EDSECPER and EDTERPER are also obtained by summing across the corresponding five year cohorts, i.e., EdPriPopPer etc.

(Note: AGEDST is in all caps in other places)

Cohort attainments by level of education are also used in to build a specialized educational attainment display, commonly referred to as education pyramid in congruence with demographic pyramids used to display population by age cohorts stacked one on top of the other with the men and women cohorts put opposite to each other around a vertical axis. Education pyramid superimposes educational attainment on top of the demographic pyramid.

# Reference

Barro, Robert and Jong-Wha Lee, "A New Data Set of Educational Attainment in the World, 1950-2010." forthcoming, Journal of Development Economics. http://www.barrolee.com/data/dataexp.htm

Barro, Robert J., and Jong-Wha Lee. 2000. “International Data on Educational Attainment: Updates and Implications.” NBER Working Paper 7911. National Bureau of Economic Research, Cambridge, MA.

Bruns, Barbara, Alain Mingat, and Ramahatra Rakotomalala. 2003. Achieving Universal Primary Education by 2015: A Chance for Every Child. Washington, DC: World Bank.

Clemens, Michael A. 2004. The Long Walk to School: International education goals in historical perspective. Econ WPA, March. http://ideas.repec.org/p/wpa/wuwpdc/0403007.html.

Cohen, Daniel, and Marcelo Soto. 2001. “Growth and Human Capital: Good Data, Good Results.” Technical Paper 179. Paris: OECD.

Cuaresma, Jesus Crespo, and Wolfgang Lutz. 2007 (April). “Human Capital, Age Structure and Economic Growth: Evidence from a New Dataset.” Interim Report IR-07-011. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Delamonica, Enrique, Santosh Mehrotra, and Jan Vandemoortele. 2001 (August). “Is EFA Affordable? Estimating the Global Minimum Cost of ‘Education for All’”. Innocenti Working Paper No. 87. Florence: UNICEF Innocenti Research Centre. <http://ideas.repec.org/p/wpa/wuwpdc/0403007.html>.

Dickson, Janet R., Barry B. Hughes, and Mohammod T. Irfan. 2010. Advancing Global Education. Vol 2, Patterns of Potential Human Progress series. Boulder, CO, and New Delhi, India: Paradigm Publishers and Oxford University Press. http://www.ifs.du.edu/documents

Hughes, Barry B. 2004b (March). “International Futures (IFs): An Overview of Structural Design.” Pardee Center for International Futures Working Paper, Denver, CO. http://www.ifs.du.edu/documents/reports.aspx

Hughes, Barry B. and Evan E. Hillebrand. 2006. *Exploring and Shaping International Futures*. Boulder, Co: Paradigm Publishers.

Hughes, Barry B. with Anwar Hossain and Mohammod T. Irfan. 2004 (May). “The Structure of IFs.” Pardee Center for International Futures Working Paper, Denver, CO. http://www.ifs.du.edu/documents/reports.aspx.

Irfan, Mohammod T. 2008. “A Global Education Transition: Computer Simulation of Alternative Paths in Universal Basic Education,” Ph.D. dissertation presented to the Josef Korbel School of International Studies, University of Denver, Denver, Colorado. http://www.ifs.du.edu/documents/reports.aspx.

McMahon, Walter W. 1999 (first published in paperback in 2002). Education and Development: Measuring the Social Benefits. Oxford: Oxford University Press.

Wils, Annababette and Raymond O'Connor. 2003. “The causes and dynamics of the global education transition.” AED Working Paper. Washington, DC: Academy for Educational Development

1. The multi-state demographic method developed and utilized by IIASA does include education-specific mortality rates. [↑](#footnote-ref-1)